VRIJE UNIVERSITEIT AMSTERDAM

VU Research Portal

Railway Impacts on Real Estate Prices

Debrezion Andom, G.

2006

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Debrezion Andom, G. (2006). *Railway Impacts on Real Estate Prices*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam]. Thela Thesis/TI.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 04. May. 2023

Railway Impacts on Real Estate Prices

ISBN 90 5170 968 4

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **389** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

Railway Impacts on Real Estate Prices

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Economische Wetenschappen en Bedrijfskunde
op dinsdag 28 november 2006 om 13.45 uur
in het auditorium van de universiteit,
De Boelelaan 1105

door

Ghebreegziabiher Debrezion Andom

geboren te Halibo, Eritrea

promotor:	prof.dr. P. Rietveld
copromotor:	dr. E. Pels

For my

Wife Martha Yitbarek

and

Son Solomon G. Debrezion

Preface

For the Last four years I have been working on my Ph.D. study based on a real estate project funded by the Netherlands Organization for Scientific Research (NWO). With the help of the ever gracious God and the support of many, now I have come to the end of writing the dissertation. Let me take the time and the space to thank the people who helped me during the course of my research.

My first words of thanks go to my supervisor prof.dr. Piet Rietveld. In the midst of many difficulties at the start of my studies, by steadfastly committing himself, he has helped me to set afoot. I thank him for this and showing confidence in me all along. During the course of the research I have gained invaluable research insights and direction from him. My questions for direction were welcomed at all times, even at the tightest imaginable schedule. Equally, I would like to thank my co-promoter dr. Eric Pels. More than anyone else, he was there for me for corrections and insights. His detailed comments were instrumental in my research works. Because of his consistency in correction, he has helped me to shape my writing style.

As I have pointed out earlier, the real estate project on which my Ph.D. study is based was funded by the Netherlands Organization for Scientific Research (NWO). This Ph.D. project also partly took place in the context of the Habiforum programme on Innovative Land Use, in particular the theme of land use around transport nodes. I would like to extend my thanks to the organization for all the financial support. Without its generous support the study would not have been possible. In addition to the financial support, the research has required different kinds of data. Acquiring the relevant data was also possible by the cooperation of many parties. My first words of thanks go to the Dutch Brokers Association (NVM). They have allowed me to use a rich and quality data set on house sales transactions for an extended period of time. Second, I have received unreserved use of many rich and quality data sets from the Dutch Railway Company (NS). Many thanks for their cooperation. Special thanks to Mr. Mark van Hagen from NS for his cooperation in this regard. Thirdly, I would like to thank DTZ Zadelhoff for providing me with data on the rent of office spaces. In that regard special thanks to Mr. Bart Louw. Lastly, but surely not least, I would like to extend my words of thanks to people in the SPINlab of the department of Spatial Economics for allowing me to use data that involved GIS and for related technical assistance. Special words of thanks go to Jasper Dekkers. He has given me unreserved cooperation in this regard. Many thanks. In

addition, I would like to thank Alfred Wagtendonk and Nancy Omtzigt for their technical assistance at the beginning of my GIS exercise.

Besides the help of my supervisors, I was turning to my colleagues for insights and reflections both at the department of Spatial Economics and Tinbergen Institute Amsterdam. Many thanks to dr. Thomas de Graaff for his insights in Spatial Econometrics, Jaap de Vries for his insights in Spatial Interaction Modelling, dr. Martijn Brons for his insight in Meta analysis and others. Literally, every one in the department has contributed to my experience in one way or another making my stay at the Tinbergen Institute Amsterdam and the Department of Spatial Economics enjoyable. Special words of appreciation go to my office mates Yin Yen Tseng, Vanessa Eve Daniel, and Olaf Jonkeren. Jośe Melendeze-Hidalgo, Romy Zang and Alex Halsema deserve words of thanks for their friendship. In addition, I thank Alex Halsema, dr. Eric Pels and Olaf Jonkeren for the translation work of the Dutch summary.

I also am very much indebted to the members of the reading committee of this dissertation: prof.dr. Luca Bertolini, prof.dr. Peter Nijkamp, prof.dr. Hugo Premius, prof.dr. Bert van Wee, and prof.dr. Mark Wardman. Many thanks for their willingness and time for reading and commenting on the draft of the dissertation.

Mrs Patricia Elleman deserves special thanks for her language correction of the dissertation. She has gone over it several times with great detail. Her corrections were very useful.

Outside the scientific circle, I also have been supported by many. My beloved wife Martha Yitbarek takes the star thanks. She has contributed a lot in making the concluding work of my dissertation without stress. And this was despite her uncomfortable months of early pregnancy for our first baby Solomon. Million thanks Martha and I love you. I also extend by thanks to my parents for supporting me all along and for preparing me to this stage from my early age. I thank my sisters and brothers for their encouragement.

Praises be to God.

Many thanks

Ghebreegziabiher Debrezion

Amsterdam, October 2006.

CONTENTS

Tables:

Figures:

Chapter 1

1 Introduction

1.1 SETTING THE SCENE

Transportation is defined as a way of moving from one place to another. It is an important activity that shapes the overall economic and social dynamics. It affects where we live, where we work, where we shop, and even what is supplied to markets. Literally every aspect of our activities is influenced by transportation in one way or the other. Clark (1958) argues that transportation was a prime factor in the rise and fall of empires in ancient history and the maker and breaker of cities in modern times. Since the invention of the wheel, which is believed to be the most important mechanical invention of all times, transportation has advanced tremendously; and so has the influence of transport. It involves advances in capacity, cost, speed, comfort, privacy, taste, etc. across different modes of transport.

Railway transport is generally considered as a symbol of the Industrial Revolution. It was instrumental in boosting production processes. The sector that enjoyed the immediate benefit of railway transport was the agricultural sector. Rail transport opened the possibility of specialized mass production in the agricultural sector. In addition to the effects on the agricultural sector, the rapid development of heavy industries was observed because of railway transport (Clark 1958). In the early stages, passenger transport mainly relied on horses, and the application of the railway was rather limited to the transportation of goods. At a later stage of the introduction of steam engine, passenger transport by rail also started to develop, though it still remained limited to long distance transport. Transport within cities was generally limited to walking, horse-drawn carriages, etc., and thus cities remained compact and limited in size (Anas et al. 1998; Oosterhaven and Rietveld 2005). Thus, the influence of railway transport on urban dynamics remained limited. But, this was only until the introduction of the electric railway, after which the railway started to revolutionize urban transport and shape urban dynamics. As a result, urban areas started to lose their compact nature, becoming more de-concentrated. Firms moved to satellite locations outside the city core to take advantage of the lower factor costs, while remaining close to railway stations to keep their link to the central core. Railway stations thus reduced the demand friction around the central business district (CBD) to some degree, by attracting commercial entities and households to locate themselves around the stations (Fejarang 1994). The de-concentration was further enhanced by buses and cars. As a result we then see transport nodes attracting development activities. In the initial stages, the success of newly opened railway stations in attracting settlement and commercial development activities mainly depended on the railway stations' relative proximity to the city's central core. As these developments continued, the nodes became more independent from the central core and grew into sub-centres. As a result, polycentric urban structures evolved.

These days, the automobile is becoming the dominant means of passenger transport; this comes at the expense of the decreasing popularity of railway transport. This has further resulted in low density, extended city size, and urban sprawl. Reviving railway transport is viewed as a viable solution to keep the integral part of the urban area from further sprawl (Goldberg 1981). Thus, the railway has to regain its competitive position. In that respect, we see several cities adopting light, heavy, and commuter rail systems. In cities such as London, Paris and New York public transport, and especially rail transport, plays a vital role. However, in many other cases, the impact of railways on the urban dynamics still mainly depends, among other things, on coordination with land use and government policies. To further increase the competitive position of railway transport compared with car, high speed train (HST) services are arriving on the scene. As will be outlined below, the effect of railway developments on property values is an important way by which the railway shapes urban forms.

1.2 TRANSPORT INFRASTRUCTURE AND LAND PRICES: MOTIVATION

Up to the beginning of the 20th century the discussion about land rent was basically about agricultural land. This was partly because urban areas were considered unproductive (Smith 1776). David Ricardo's (1821) seminal work on agricultural land rent indicates that land rent is determined by the fertility level of the agricultural land. The difference in the fertility level of agricultural land is reflected in the differences in land rent. He further recognized that proximity to the market is capitalized in the land rent, though no deeper investigation was made. In a subsequent study, Von Thünen (1863) investigated the effect of proximity to the market on agricultural land rent. For a given fertility of land, land rent declines with distance from the isolated city's centre. Furthermore, the model highlights the land use patterns of the agricultural land. The immediate agricultural land area is devoted to products, such as diary products, vegetables and fruits that require rapid access to the market. The outer ring of the agricultural area is devoted to products that do not require quick access to the market, or otherwise involve self-transport such as ranching.

Since the work of Von Thünen, economists have addressed the issue of the relationship between the impacts of the improvement of transportation infrastructure and land prices in the urban context. Economists like Alonso and Muth refined Von Thünen's line of reasoning into a bid-rent analysis (Alonso 1964; Muth 1969). The basic idea behind the bid-rent model is that every agent is prepared to pay a certain amount of money, depending on the location of the land. This leads to a rent gradient that declines with distance from the CBD because of the increase in transportation cost for sites that yield equal utility. Thus far, the urban models assumed a monocentric city which described the city as a circular residential area surrounding a CBD in which all jobs are located, served by a radial transport system. For several decades, the monocentric model remained popular in explaining urban spatial structures. However, additional complications to the urban model started to arise as a result of the spatial structure of urban areas. As urban areas grow more and more polycentric in nature, classic urban models based on monocentric assumptions are coming under pressure to accommodate these changes. Polycentricity denotes the existence of multiple centres in a region (Davoudi 2002). In these cases, the influence of urban centres on the bid-rent function of land is not only limited to the historic CBD. Land rent is also influenced by proximity to local centres. The influence of urban centres on the bid-rent functions of urban land depends on the importance of the centres as destination points for the economic activities of households. This situation calls for a shift in the urban modelling exercise from a monocentric city assumption into a polycentric city assumption.

One of the main reasons for the development of urban areas into a polycentric structure is the increase in mobility as a result of new transport technologies. The increasing use of private cars is believed to be instrumental in shaping the present dispersed urban structure (Clark 1958; Glaeser and Kahn 2004). As a natural response to reverse the growing congestion posed by automobile traffic and urban decentralization, the development of the railway is starting to revive in many parts of the world. Railway investment is expected to support a more compact urban structure and therefore it serves the urban planning purpose (Goldberg 1981). Thus, the importance of both modes of transport (i.e. car and rail) in shaping the urban structure also has a modelling implication. So far, the classical monocentric urban models have assumed that a radial unimodal transport system operates in the urban area. Thus, this consideration will lead us to the adaptation of the urban model to include a multimodal transport system in the urban settings.

Moreover, according the model of Von Thünen (1863), it is the landowner who finally absorbs the benefits of a uniform improvement of infrastructure. In more complex urban land use models, the total welfare effects of an improvement of infrastructure are shared equally between landowners and residents (Fujita 1989). Most theoretical results concerning the impacts of infrastructure improvements on land prices are based on the assumption that the infrastructure improvement is uniform (for example, a uniform increase in speeds). However, as indicated by Mohring (1993), in the case of a non-uniform improvement of infrastructure (for example, the construction of a highway in addition to a low speed network), the effects on land prices may be quite different. Even more differentiated effects may be expected in the case of a multimodal transportation network with transfer nodes such as railway stations. This calls for an analysis of land prices in an urban system with multiple transport nodes, including railway stations.

In order to understand the contribution of railway stations to the dynamics of urban areas, it is necessary to understand the effects that railway stations will have on prices, since prices are important signals to developers. Of particular importance is the problem of mobilizing sufficient resources for the construction of railway lines. The potential for the development of real estate around railway stations can be assessed by means of the models developed here. Hence, it is possible to find out to what extent the costs of building railway lines and railway stations can be covered by means of the participation of real estate developers. Later in this thesis the case study of the Amsterdam South Axis development is used for forecasting the impact of HSL South both on residential and commercial properties. The development concerns the largest infrastructure-related urban development project in the Netherlands.

1.3 RESEARCH QUESTIONS

Taking the modelling and empirical considerations discussed in the introduction, this thesis aims to investigate the effects of the construction of railway stations on land prices in urban systems. In addition to a theoretical formulation of the relationship between railway developments and land prices, empirical models will be estimated. The central question addressed in this thesis is:

> *How are land prices (for housing and offices) affected by the development of multimodal transport nodes in general, and railway stations in particular?*

At different points in the thesis we further address a number of sub-questions. These sub-questions are aimed at presenting an appropriate approach to fully address the central question of the thesis stated above. As we discuss the organization of the thesis, we will indicate which of the following sub-questions is addressed in which part of the thesis.

1. What can be learned from existing empirical studies on the area?
2. What are the implications of regulation on the land market for the effect of railway development on land prices?
3. What is railway accessibility? How can it be made operational in impact analysis?
4. What is the role of the access mode to a departure station on overall railway accessibility?
5. What is, in empirical terms, the contribution of railway accessibility to the explanation of prices of offices and residential dwellings?
6. What are the implications of HSL (high speed line) South for the Amsterdam South Axis Station and its effect on residential and office prices?

1.4 METHODOLOGICAL APPROACH

In order to answer the above-mentioned main research questions of this thesis, several methodological approaches are followed. Below, we briefly discuss the approaches pursued in this thesis:

a) ***Meta-analysis***: as a starting point, after a brief discussion of the theoretical foundations of the area, the thesis undertakes an intensive literature review of empirical research outcomes concerning the effect of railway stations on property values. In addition to the qualitative review of the studies, the thesis conducts a quantitative analysis to explain the difference in the results of the different empirical studies. This is based on a meta-analytical approach. Meta-analysis is a statistical analysis that combines the results of independent studies in an effort to explain the differences in study results by study settings. This methodology is most popular in the medical sciences. However, it is increasingly used in economic analysis.

b) ***Modelling and simulation analysis***: The model will be formulated in line with the approaches proposed by Miyao (1981), Fujita (1989), Fujita et al. (1999) and Medda (2000). The model can be used to analyse the effects on land use and land prices of developments in railway lines and networks. The introduction of multimodal transport nodes adds a special element to the theory that is mainly based on monocentricity. It may also provide an interpretation of the phenomenon of 'edge cities' (Garreau 1988). The land market imperfections may relate to the external effects of one type of land use on the other. Government interventions such as building restrictions at certain places and parking policies can also be analysed in this context. These imperfections and policies may have strong effects on urban development and on urban land prices (Rietveld et al. 2001). In this thesis, we look at the effects of the commercial land restrictions. A comparison is made concerning the implications of underlying land markets for the effect of railway investments. The thesis formulates a theoretical urban land use model with multimodal transport nodes according to the theory of urban economics, taking into account the implications of different land market regimes. The model is used to investigate the consequences of various types of changes in multimodal transport networks on land prices near railway areas.

c) ***Measuring railway accessibility:*** One of the trickiest aspects of in this type of studies is how to measure accessibility properly. In the literature, several definitions and ways of measuring accessibility exist. In this thesis, we define railway accessibility in terms of the ease of reaching a railway station and the level of rail and supplementary services provided at the railway station. An overall rail service measure of a particular railway station in the network is determined through spatial interaction model analysis. Furthermore, this thesis acknowledges the drawback characteristic of most empirical studies in this area concerning the measurement of general railway accessibility. Mostly, accessibility to railway stations is discussed in connection with the nearest railway station. However, in this thesis we noted real estate prices react to more factors than just the closest railway station. Travellers mostly have a set of railway stations which they use as departure stations. This phenomenon is explained by carrying out a choice analysis on departure stations. We assume that the accessibility of a location (a house, an office, etc.) to railway transport is explained by a number of factors related to the ease of reaching the railway station in an individual's choice set and to the rail and supplementary services provided at the railway stations. Different access modes can be used to reach the station. The general railway accessibility is therefore an aggregate function of these features over the entire set of railway stations in the choice, weighted according their degree of importance. Thus, based on both access mode and departure railway station choices, a nested logit model is estimated with the ultimate aim of computing the general railway accessibility at a location.

d) ***Hedonic pricing estimation analysis***: The model that will be estimated is essentially a hedonic price model (see, e.g. Rosen 1974), where prices of housing are explained by the internal properties of houses, accessibility, and in particular by their location with respect to railway stations, and the quality of the services supplied via these transport nodes. Data will be used on the prices of dwellings in the owner-occupier part of the housing market (NVM data), which will be linked to spatial data via GIS to take the neighbourhood effects into account. In addition, linkage will be made using specific transport network data in order to take into account the quality of multimodal transport networks, and in particular the role of railway stations. Along similar lines, a model will be estimated for office rents (for an earlier study of office rents in the Amsterdam area and the impact of railway locations, see Rietveld and Bruinsma 1998, Chapter 9). We estimate an empirical model of the prices of housing and offices, where special attention is paid to the impact of multimodal accessibility via railway stations.

e) ***Spatial autocorrelation analysis:*** The issue of *spatial autocorrelation* (Anselin and Florax, 1995) is addressed in the thesis. Spatial autocorrelation means that the outcomes of processes that are located in close proximity may be correlated because of unobserved neighbourhood effects. This is an important issue, because the unobserved neighbourhood effects may interfere with the analysis of the impact of the location with respect to railway stations. The relevance of spatial autocorrelation in hedonic price models has recently become a point of attention in research (see, for example, van der Kruk 2001).

f) ***Application of estimated model:*** This is done by carrying out ex-ante assessments of the effects of the creation of railway lines and railway stations. This entails the use of the model for a specific infrastructure project in order to assess the impacts on changes in real estate prices. An interesting area of application will be the Amsterdam South Axis (the Zuidas) that will eventually be a station for two international high-speed railway lines and where the local infrastructure has recently been improved, while further improvements are underway (the North-South line).

1.5 RESEARCH SCOPE

Investment in transportation infrastructure has a wide range of economic effects. The impact focus can take different forms: either a broad macro-perspective, such as an analysis of the impact of the transport infrastructure on the economy of a certain geographical area, employment etc., or a micro-perspective, such as an analysis of the impact of transport infrastructure on property values. However, talking in terms of transport infrastructure can still be quite broad in that transport infrastructure encompasses different modes of transport. In many theoretical and empirical analyses, it has been indicated that prices of real estate tend to react to changes in transportation infrastructure. Even though car-based transportation dominates, railway transportation also has a non-negligible share in most European cities. This thesis investigates the effect of railway stations on the prices of real estate. Its focus will be on urban land, and built-up areas of offices and residences.

Several novel contributions to the literature in the area can be found in this thesis. In the previous literature, several review studies have been conducted. However, in this thesis we present a statistical analysis of the empirical results reported in the literature. To our knowledge this effort is the first meta-analytical study conducted in the area. Moreover, the thesis develops a polycentric urban model in the context of multimodal transport. Even though the polycentric urban model is not new, giving it a multimodal dimension is a new effort here. Furthermore, our methodology of measuring railway accessibility by taking into account modal choice in accessing the railway station is fairly unique. Finally, our approach of taking into account spatial autocorrelation in the analyses of railway impacts is novel as far as we know.

1.6 RESEARCH ORGANIZATION

The remainder of the thesis is organized in eight chapters. In Chapter 2 we review the theoretical literature. Special attention will be given to hedonic pricing theory. This will be followed by a review of the empirical literature in the area concerning the impact of transport infrastructure on the prices of real estate. On the basis of this review, we make an analysis of the literature findings using meta-analytical methodology. In Chapter 3, the thesis addresses modelling issues. Departing from the classical monocentric model, it discusses the polycentric multimodal transport urban model. Equilibrium conditions for a selected parameter set will be determined. The analysis and discussion is based on the bimodal bicentric open city case. In addition, we discuss the implications of different land markets for the impact of railway investment on the prices of real estate. The analysis is supported by simulations. In Chapter 4, baseline hedonic price estimation of the effect of railway accessibility on residential house prices is discussed.

The next two chapters play an instrumental part in the design of the thesis. They deal with ways of measuring railway accessibility. Chapter 5 discusses a way of addressing the rail service quality of a railway station, which is expected to influence the real estate value. Based on a spatial interaction model, a rail service quality index (RSQI) of all railway stations in the Dutch railway network is determined. In a bid to determine the general railway accessibility of a location, the thesis undertakes a choice analysis for a departure railway station and accompanying access mode. This is discussed in Chapter 6. It builds on the previous chapter by including the access part of the trip in determining the overall accessibility in relation to railway transport. The outcomes of both chapters are used in the hedonic price model estimation of real estate value.

Chapters 7 and 8 discuss, respectively, the estimations concerning the residential and commercial property market of the Netherlands. In Chapter 7 a spatial autocorrelation model is estimated for the residential property value. It includes the overall railway accessibility measure determined in Chapters 5 and 6. In Chapter 8 we estimate a spatial autocorrelation model for office space rent levels. In both chapters, the implications of High Speed Line (HSL) South for the South Axis with respect to residential property value and office rental levels are discussed.

Finally last Chapter 9 gives the summaries and conclusion of the thesis.

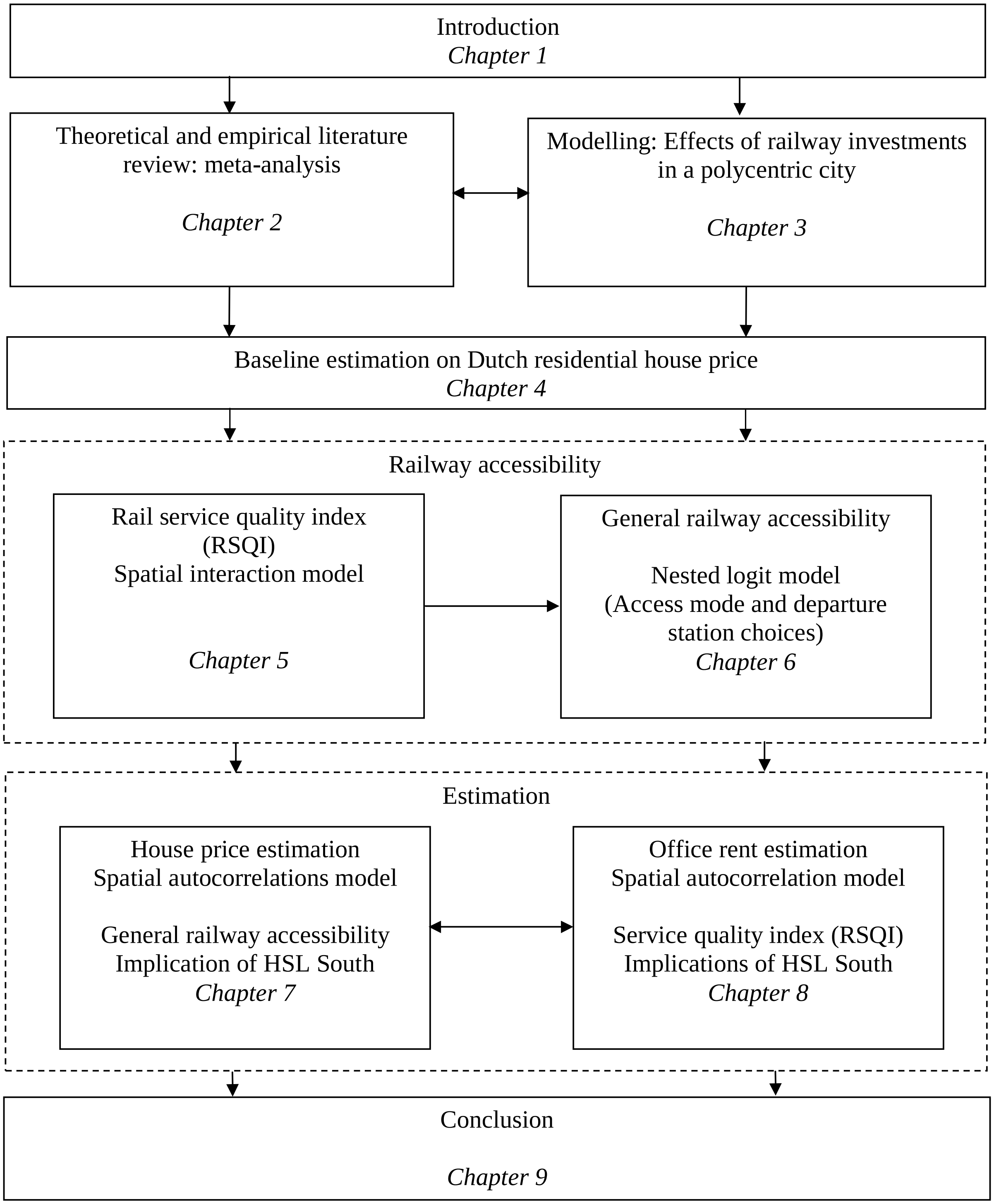


Figure 1.1: Thesis layout

Chapter 2

2 The impact of railway stations on residential and commercial property value: a meta-analysis[1]

2.1 INTRODUCTION

Location choice is a frequently discussed topic in urban economics. These discussions can be normative or descriptive in nature. In the literature, we find two approaches to urban location analysis. The first set of studies addresses the issue of optimal location conditional to a given set of constraints (Fujita 1989). The second set of studies is devoted to explaining the character (value) of a property at a given location. However, the issue of identifying the factors that affect property values is common to both sets of approaches. Our discussion in this thesis basically addresses studies of the latter category, focusing on the relationship between property values and railway stations. In the context of this thesis, property means an estate ranging from a vacant piece of land to an area occupied by all sorts of buildings: residential, commercial, industrial, etc. (Brigham 1965). Several studies have tried to address the various discussions on property values. There is a general consensus among most authors in categorizing the factors affecting property values as physical, environmental, and accessibility factors (Fujita 1989; Bowes and Ihlanfeldt 2001). However, some authors have included historical factors and land use patterns in their analysis (Brigham 1965). Numerous detailed lists of features can be identified within each of these categories. As to the relevance of the factors to the analysis, the detailed list can differ from one place to another, and thereby from study to study.

Accessibility, as provided by different modes of transportation and the railway in particular, as a factor affecting property values, has also received some attention in the literature. The most common way of addressing railway accessibility has been by including the proximity factor in the analysis. This chapter discusses the results of studies which have addressed the

[1] This chapter is based on Debrezion, Ghebreegziabiher, Eric Pels and Piet Rietveld (2006). “The Impact of Railway Stations on Residential and Commercial Property Value: A Meta Analysis.” *Journal of Real Estate Finance and Economics*, Forthcoming.

effect of railway stations on property values. The chapter has two parts: a qualitative review, and a quantitative analysis. The first part of the chapter surveys studies on the effect of railway station proximity on property values. In the second part, meta-analytical analysis is applied to systematically explain the variation in the findings on the impact of railway station proximity on property values across studies. Thus, in the subsequent sections, we discuss the theoretical foundation of the studies, presenting and comparing the empirical results of the various studies that have been undertaken. In addition to the review of studies conducted in the area, we make a quantitative analysis of the results of the studies, using meta-analysis to explain the differences in the results.

2.2 LITERATURE REVIEW: THEORIES AND EMPIRICAL FINDINGS

Most land value theories have their roots in the work of Von Thünen (1863), who tried to explain variations in farmland values. According to Von Thünen, accessibility to the marketplace explains the difference in value of areas of agricultural land with similar fertility. In subsequent studies, economists like Alonso and Muth refined this line of reasoning into a bid-rent analysis (Alonso 1964; Muth 1969). The basic idea behind the bid-rent model is that every agent is prepared to pay a certain amount of money, depending on the location of the land. This leads to a rent gradient that declines with distance from the central business district (CBD) for sites that yield equal utility. Thus far in the analyses, the dominant factor explaining the difference between land (property) values was the accessibility as measured by the distance to the CBD and the associated transportation costs. The physical characteristics of the land (fertility in the case of Von Thünen) were assumed given.

Thus, the basic theory on real estate prices can be explained as follows: as a location becomes more attractive, as a result of having certain characteristics, demand increases and thus the bidding process pushes prices up. In most cases CBDs are the centres of many activities. Therefore, closeness to the CBD is considered as an attractive quality that increases property prices. However, investments in transport infrastructure reduce this demand friction around the CBD itself to some degree (Fejarang 1994) by attracting households to settle around the stations with rapid access to the CBD. Properties close to the investment area (railway stations) enjoy benefits from transportation time and cost saving as a result of the investment.

It may be expected that a price curve will have a negative slope: with increasing distance away from the station, prices decrease.

The introduction of the hedonic pricing methodology (Rosen 1974) led to an easier way of attributing effects on property values to various features of the properties. Thus, we observe the integration of physical, accessibility, and environmental characteristics of properties in models trying to explain the difference in property values. Accessibility remains an important feature for urban properties. However, earlier attempts to account for it by transportation cost have been too narrow. Attempts have been made to introduce a broader concept of accessibility so as to include all features that contribute to a location's potential for interaction (Hansen 1959; Martellato et al. 1998). Though a comprehensive definition of the concept of accessibility is available, in practical applications the lack of data and appropriate measuring techniques have usually implied that only simple measures have been used. Thus, in the literature we see a focus just on some factors, especially a CBD-oriented interaction related to employment and shopping. In most property value studies, the other trip purposes (e.g. entertainment, leisure, etc.) are missing from the model.

The main focus in this chapter is the analysis of the impact of railway accessibility as measured by proximity to railway stations. However, it is important to realize that accessibility can also be provided by other modes of transport. As Voith (1993) has pointed out, highway accessibility is an important competitor to rail accessibility: "The presence of other facilities that increase accessibility like highways, sewer services and other facilities influence the impact area in the same fashion". The benefits of these facilities and services are also capitalized in urban property values (Damm et al. 1980). Thus, to single out the effect of railway accessibility, other competing modes of accessibility need to be included along with it.

The motivations for the studies on the impact of railway accessibility are diverse. The larger part of the literature on the railway focusses on it as a feasible solution to the rising congestion posed by automobile traffic and urban sprawl. Railway investment is expected to support a more compact urban structure, and therefore it serves the urban planning purpose (Goldberg 1981). Apart from attempting to show that railway investments do result in compact urbanization, most studies in the area were conducted to provide evidence for the implementation of value capture schemes for financing rail investments (Cervero and

Susantono 1999). This was based on the assertion that the value of proximity to accessibility points is capitalized in the value of properties around these stations.

In general, the empirical studies conducted in this area, i.e. on the impact of railway accessibility on property values, are diverse in methodology and focus. Although the functional forms can differ from study to study, the most common methodology encountered in the literature is hedonic pricing. However, no consistent relationship between proximity to railway stations and property values is recorded. Furthermore, the magnitude of these effects can be minor or major. In one of the earliest studies, Dewees (1976) analysed the relationship between railway travel costs and residential property values. Dewees found that a subway station increases the site rent perpendicular to the facility within $\frac{1}{3}$ of a mile from the station. Similar findings confirmed that the distance of a plot of land from the nearest station has a statistically significant effect on the property value of the land (Damm et al. 1980). Consistent with these conclusions, Grass (1992) later found a direct relationship between the distance of a newly opened metro and residential property values. Some of the extensively studied metro stations in the U.S., though ranging from small to modest in impact, show that properties close to the station have a higher value than properties farther away (Giuliano 1986; Bajic 1983; Voith 1991). However, there are also studies which have found insignificant effects (Lee 1973; Gatzlaff and Smith 1993). On the other hand, contrary to the general assumption, Dornbusch (1975), Burkhart (1976) and Landis et al. (1995) traced a negative effect of station proximity. Evidence from other studies indicates little impact in the absence of favourable factors (Gordon and Richardson 1989, Guiliano 1986). For a detailed documentation of the findings, we refer to Vessali (1996), Smith (2001), NEORail II (2001), Hack (2002), and RICS (2002). In general, some studies indicate a decline in the historical impact of railway stations on property values. This was attributed to improvements in accessibility, advances in telecommunications, computer networks, and other areas of technology that were said to make companies “footloose” in their location choices (Gatzlaff and Smith 1993).

Our main aim in this chapter is to systematically analyse the variation in the findings of the studies discussed above. We use meta-analysis to provide a statistical analysis of the variations in the study findings. The impact of a railway station on the property values depends on several factors.

First, railway stations differ from each other in terms of service levels provided such as frequency, network connectivity, service coverage, etc. Thus, it is natural to see stations with differing impacts on the value of surrounding properties. Commuter railways have a relatively high impact on property values (Cervero and Duncan 2001; NEORail II 2001; Cervero 1984). Railway stations can also differ in the level and quality of facilities they have. Stations with a higher level and quality of facilities are expected to have greater impact on the value of surrounding properties. The presence and number of parking lots is one of the many station facilities that have received attention in this area. Bowes and Ihlanfeldt (2001) found that stations with parking facilities have a higher positive impact on property values. In addition, the impact that a railway station produces depends on its proximity to the CBD. Stations which lie close to the CBD produce a greater positive impact on property values (Bowes and Ihlanfeldt 2001). In addition Gatzlaff and Smith (1993) claimed that the variation in the findings of the empirical work is attributed to local factors in each city.

Second, railway stations affect residential and commercial properties differently. Most studies have treated the effect of railways on the different property types separately. That allows us to explain the difference of railway effects on different property types. In general, it has been shown that, within short distance of the stations, the impact of railway stations is greater on commercial properties compared with residential ones. The greater part of the empirical literature on property value focusses on residential properties rather than commercial properties. Usually, it is claimed that the range of the impact area of railway stations is larger for residential properties, whereas the impact of a railway station on commercial properties is limited to the immediately adjacent areas. But there are also claims that railway stations have a higher effect on commercial than on residential properties (Weinstein and Clower 1999; Cervero and Duncan 2001). This finding is in line with the assertion that railway stations as focal, gathering points attract commercial activities, which increase commercial property values. However, contrary to this assertion, Landis et al. (1995) determined a negative effect on commercial property values.

Third, the impact of railway stations on property values depends on demographic factors. Income and social (racial) divisions are common. Proximity to a railway station is of higher value to low-income residential neighbourhoods than to high-income residential neighbourhoods (Nelson 1998; Bowes and Ihlanfeldt 2001). The reason is that low-income residents tend to rely on public transport and thus attach higher value to living close to the station. Because reaching the railway station mostly depends on slow modes (walking and bicycle), the immediate locations are expected to have higher effects than locations further away. On the other hand, the high population movement in the immediate location gives rise to the development of retail activities which leads to premiums on the value of commercial properties. But, at the same time, these retail properties may attract criminality (Bowes and Ihlanfeldt 2001). Bowes and Ihlanfeldt (2001) observed that a significant relation between stations and crime rates. In their model, the immediate neighbourhood is negatively affected by the station. Thus, the most immediate properties (within ¼ of a mile of the station) were found to have an 18.7% lower value. Properties that are situated between one and three miles from the station are, however, more valuable than those further away. Though this study provides an important contribution, unexplained variations still remain.

2.3 META-ANALYSIS OF THE STUDIES

In the previous section, we briefly reviewed empirical work on the effects of station proximity on property value. Other reviews can be found in Vessali (1996), Smith 2001, NEORail II (2001), Hack (2002), and RICS (2002). These studies also summarized empirical work in this area, but did not look for a systematic explanation of the variation in the findings. Our study not only summarizes earlier work, but also looks for a systematic explanation of differences in the results. Meta-analysis serves as an important tool for this purpose (Smith and Huang 1995; Cook et al. 1992). It provides statistical synthesis for empirical research focussed on a common research question. It includes variables that represent study settings that are expected to explain the variation in the findings of the studies. In this case, all the reviewed studies focus on the impact of railway station proximity on property values. For the comparison of results to be meaningful, it is required that the studies have a comparable unit for the effect. However, in the studies which address the relationship between proximity to a railway station and property values, we encounter different measurement units, although they aim at measuring similar effects. Thus, it is important that the findings are converted into the same measurement unit.

In this study we apply a meta-regression model. The effect sizes of proximity to the railway station on property values found by the different studies are the dependent variables, whereas the implicit or explicit characteristics of the underlying studies make up the independent variables. A basic meta-analysis equation can be given as follows (Florax et al. 2002).

$$Y = f(P, X, R, T, L) + \varepsilon , \tag{1}$$

where *Y*= the variable under study;

P= the set of causes of the outcome Y;

X= the characteristics of the set of objects under examination affected by P in order to determine the outcome Y;

R= the characteristics of the research method;

T= the time period covered by the study;

L= the location of each study conducted;

ε = the error term.

2.3.1 Model specification

Meta-analysis models try to explain the difference in study findings by difference in study characteristics and other variables: for instance, time and geographical effects. Thus, generally they belong to the family of hedonic pricing models. The logical order is first to identify the characteristics of the underlying studies that could explain the variations in effect sizes. The underlying studies usually include the proximity of the property to the station. However, we observe that not all studies use the same set of (explanatory) variables. The studies also differ in methodology. A railway station variable is mostly treated as a sole indicator of the accessibility of a certain area. However, other modes serve the same purpose; for example, highway/freeway presence in the area under consideration. Although for our purposes, it is important to note that they both have an effect on property values, it is expected that these modes ‘interact’ in a complementary (one can take a car to the railway station and then take the train) or competitive way (one can use car or train).

The underlying empirical studies employ different specifications: namely, linear, semi-logarithmic, and log linear. In some studies the analyses are non-parametric in nature. Different specifications may also lead to different outcomes. In our analysis we further include type of railway station (light rail, heavy rail/Metro, commuter rail and Bus rapid transit), type of property (commercial, residential). We leave out the location feature of the studies from our model because all the studies that are used in our final analysis were done in the US. We also examine whether the underlying study includes variables for the features of the properties and demographic features. All studies include features of the property in their analysis. Thus, our analysis includes six categories of variables to explain the difference in the findings of the impact of railway station proximity on property values. To account for the variation, we specify a standard hedonic price model using a simple linear regression specification given by Equation 2 below.

$$Y = \alpha_0 + \beta_1 P + \boldsymbol{\beta}_2\mathbf{S} + \boldsymbol{\beta}_3\mathbf{M} + \beta_4 ACCESS + \beta_5 DM + \beta_6 T + \varepsilon. \quad (2)$$

Dependent variable:

Y is the effect size for the impact of railway station proximity on property values (rents) in percentages.

Explanatory variables:

P is a dummy variable that takes on the value 1 when commercial properties are analysed (residential properties are taken as the reference group). **S** is a vector of dummy variables for the station type (heavy rail/ Metro, commuter rail, bus rapid transit (BRT); light rail is the reference group). **M** is a vector of dummy variables for the model type (semi-log, double-log, non-parametric; the reference group is linear). ACCESS is a dummy variable indicating the inclusion of other means of access to the area in the underlying study (usually highways and/or freeways). DM is a dummy variable indicating the presence of a demographic variable in the underlying study (usually income or racial composition of city quarters). T is a dummy for time trend (assume 1 for study data after 1990; study data before 1990 are taken as the reference group).

Some of these variables were used in the models of the underlying studies. Others, however, relate to the settings of the studies. Because most variables in the meta-analysis are dummy variables, the estimated coefficients represent the percentage contribution of each attribute to property values in comparison with the reference groups.

2.3.2 Data and methodology

The database for the analysis of this chapter is a pool of studies concerning the impact of railway station proximity on property values. A wide range of studies is covered. A total of 73 estimation results were obtained from the underlying studies. All these studies try to quantify the impact of proximity to a railway station on property values. Different specifications in the same underlying study are treated as separate observations. Thus, the total number of underlying studies is less than the number of observations in our meta-analysis. However, because of the incompleteness of some of the studies with respect to the requirements of this study, we had to exclude certain observations. Our final estimation is based on 57 observations.

2.3.2.1 Variation in the presentation of the findings

The dependent variable in our meta-analysis is expressed as the percentage change in property values per some distance measure to the station. The underlying studies are quite diverse in the way the impact of railway station proximity is reported, including pure monetary effects, percentage effects, and elasticity measures. However, the larger part of these studies reports the percentage increase or decrease in property values for a certain distance. In addition to the diversity of measurements, the studies also use a variety of methodologies. We summarize them in two categories; which are discussed next.

I. Studies using parametric estimation methods

These studies use econometric methods to estimate the impact of railway station proximity on property values. Linear, semi-log, and log-linear (also called double-log) specifications are common. Two categories of railway station proximity measurement were encountered.

1. Station proximity as a continuous measure:

These studies consider the proximity to a railway station as a continuous variable. The variable can be measured in distance, time (walking time) or monetary savings (Dewees 1976; Nelson 1992; Benjamin and Sirmans 1996; Lewis-Workman and Brod 1997; Chen et al. 1998; Gatzlaff and Smith 1993). Sample representations of the effects of this type are given in Table 2.1. The results are given in monetary units (as in linear models) or in percentage units (as in semi-log and log-linear models). The results of the semi-log models are in line with the dependent variable in our meta-analysis. Therefore, the monetary changes and elasticities have to be transformed into a percentage change per distance using the average property value and average distance data reported in each underlying study. Coefficients of semi-log and double-log specifications represent incomparable measures. To bring them into comparable, units we divided the elasticity by the average distance of the impact area. The rent curves can have structures similar to that in panel (a) in Figure 2.1 below.

Table 2.1: Sample of railway station effects on property value based on continuous proximity measures

Author	Railway station impact on property value
Dewees (1976)	$2370 premium per hour of travel time saved for sites within 20 minutes travel time (e.g. 1/3 mile walk)
Nelson (1992)	$1.05 per foot distance to the station. premium on property value in low-income areas; $.96 per foot distance to the station.
Allen et al. (1986)	$443 premium on property value for every dollar saved in daily commuting costs (average >$4,500 per house; 7.3% of mean sales price).
Lewis-Workman and Brod (1997)	Elasticity of 0.22 with respect to property value and distance.
Benjamin and Sirmans (1996)	Rent decreased by 2.4% to 2.6% for each one-tenth mile distance from the metro station.

2. Station proximity as a distance category measures:

These studies treat the proximity variable as a discrete variable (represented by a dummy). The area under consideration is segmented into two or more parts, where the outer segment is treated as the reference (McDonald and Osuji 1995; Fejarang 1994; Dueker and Bianco 1999; Weinstein and Clower 1999; Voith 1993; Armstrong 1994; Grass 1992; Bowes and Ihlanfeldt 2001; Cervero and Duncan 2001, 2002a, 2002b; Weinberger 2001). A sample of presentations of the effects of this type is given in Table 2.2. The rent curve for these types can be given by panel (b) in Figure 2.1 below.

Table 2.2: Sample of railway station effects on property value based on distance category measures

Author	Result
Cervero, Robert (1996)	+10- 15% in rent for rental units within 1/ 4 mile of BART
Bowes and Ihlanfeldt (2001)	Property value effect (percentage change)
0-1/4 mile	-18.7%
¼-1/2 mile	2.4 %
½-1 mile	0.9 %
1-2 mile	3.5%
2-3 mile	3.5%
Weinberger (2001)	Rent
0-1/4 mile	+13 cents per square foot
¼-1/2 mile	+7 cent per square foot
½-3/4 mile	+ 1 cent per square foot
¾-1 mile	No effect

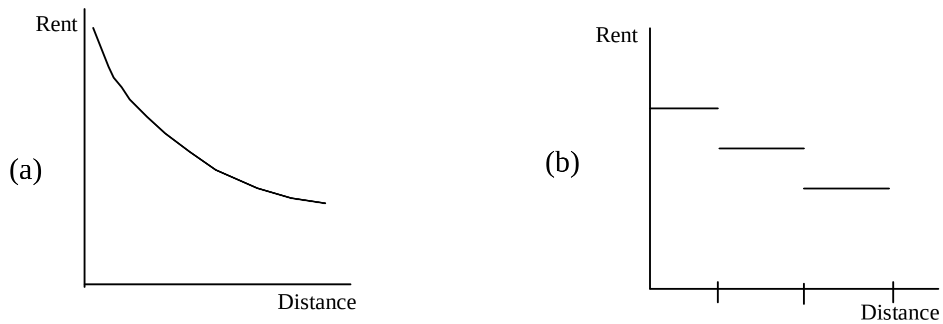


Figure 2.1: Structure of rent curves: Distance from the station as a continuous measure (a) and as category measures (b).

II. Non-parametric measures:

These studies do not use econometric methods to estimate the effect of railway stations on property values. They can measure the proximity variable in continuous or discrete terms. The common feature of these studies is that the difference in property values is implicitly attributed to the railway station effect only. Some examples of this kind of study are given in Table 2.3.

Table 2.3: Sample of results presentation for non-parametric cases

Author(s)	Result
Weinstein and Clower (1999)	Effect of station on property value
	Within ¼ mile of the station (percentage change)
Retail	36.75%
Office	13.85%
Residential	5.97%
Industrial	7.68%
Dueker and Bianco (1999)	Property value declines $1593 for every 200 feet away from the station
Fejarang (1994)	Properties within ¼ mile of the station enjoy a premium of $31 per square foot.

2.3.2.2 The dependent variable in the meta-analysis

For meta-analysis it is essential that the dependent variable is measured in comparable units. Because of the diverse ways of presenting the effect sizes, a matching process was necessary to transform them into effect sizes of the same measurement unit. For the purpose of our analysis, two proximity measuring considerations are selected: a stepwise treatment, and continuous treatment of proximity. From the standpoint of the stepwise treatment of distance, the effect of railway station proximity on values of properties located within ¼ mile of the station was prominent. Thus, we prepared the effect of railway stations on the property value for properties located within this range compared with the effect on the properties beyond this range. In addition, an effect size for the continuous distance treatment was prepared. For this consideration, the effect sizes of railway station proximity impact on property values are prepared every 250 metres closer to the station.

Because of the large differences between the underlying studies in reporting the findings, some conversion mechanism is required. We mention three elements of this mechanism:

1. We consider railway station impacts up to a maximum distance of two miles, unless otherwise indicated.
2. The properties under study are evenly distributed in concentric circles around the railway stations. Thus, because larger circles lead to an area enlargement, the average distance to the station for each segment is given by a+2/3*(b-a), where a is the distance from the border of the inner concentric circle to the station, and b is the distance from the border of the outer segment to the railway station. For the station itself we have a=b=0.
3. The impact of a station in the same segment in a circle is uniform.

For studies that provide the impact for several segments, the continuous railway station impact (see, for example, Table 2.2) is estimated by the approach outlined in Appendix 2AI. However, for studies that looked at one (inner) segment, as compared with the outer segment, we have estimated the continuous station effect per distance by point estimation (under the above assumptions). The type of model used to determine the effect can actually influence the effect (compare, for example, point elasticity estimates to interval estimates). Although most studies were parametric, a few studies used a non-parametric model, as discussed above. We adopt a unit of measurement equal to 250 metres. Thus, the dependent variable in the meta-analysis is the percentage change in property values (rents) for every 250 metres nearer to the railway station. In addition, we have prepared the effect of the railway station on the immediate segment (within a ¼ mile of the station). Therefore, our estimation is based on these two data sets.

2.3.2.3 Independent Variables

The impact of railway station proximity on property values, as reported in the underlying studies, can be affected by several factors. The type of property values under study may be important, because commercial and residential properties may be affected differently. Different types of railway stations may have different impacts because the frequency of service or the service coverage may be different, etc. Four types of rail transit services are identified: light, heavy, commuter and bus rapid transits. Three types of parametric models were encountered: linear, semi-log, and log linear. The temporal effect is represented by dividing the data into two categories: data before 1990, and data after 1990. We also included a variable for the presence of other accessibility variables (highways and freeways are of interest here), and demographic features in the underlying studies, as discussed above. As shown in Table 2.4, these considerations lead to six categories of dependent variables in our meta-analysis.

Table 2.4: Independent variables

	Variable	Description	Type
Type of property (*P*)			
	RESIDENTIAL	Residential property	Dummy
	COMMERCIAL	Commercial Property	Dummy
Type of station (**S**)			
	LRT	Light rail transit station	Dummy
	HRT	Heavy (rapid) rail transit station/ Metro	Dummy
	COMMUTER	Commuter rail transit station	Dummy
	BRT	Bus rapid transit station	Dummy
Type of underlying model (**M**)			
	LINEAR	Model with linear specification	Dummy
	SEMI-LOG	Model with semi-log specification	Dummy
	LOG-LINEAR	Model with log-linear specification	Dummy
Inclusion of accessibility variable(s) in the underlying model			
	ACCESSIBILITY (*ACCESS*)		Dummy
Inclusion of demographic variable(s) in the underlying model: income, racial composition of city quarters			
	DEMOGRAPHIC (*DM*)		Dummy
Time of data (*T*)			
	TIME Before 1990		Dummy
	TIME After 1990		Dummy

2.3.3 Descriptive statistics

Table 2.5 presents the descriptive statistics of the dependent variable. Overall characteristics and characteristics per group (defined by the independent variables) of the dependent variable are given. The overall mean impact of a railway station on property value for properties that are located within ¼ mile of the station compared with the value of properties situated beyond this range is 8.10%. The range of the property value effect is considerable: -61.90% to 145%. Concerning the continuous distance measure, the impact of a station on property values (rents) for every 250 metres closer to the station is 2.61%. The table shows that the range is considerable; it varies from –12.84% to +38.70%. In computing the means, no weighting is applied.

From Table 2.5 we also learn that railway stations have a higher average effect on commercial properties compared with residential properties. However, the corresponding standard deviations are quite high. Commuter railway stations have a higher impact on property values than the other three types of railway stations. Contrary to the assertion in the literature that railway stations have a higher impact on multi-family or condominium properties, as compared with single-family properties, the table indicates a higher impact on single-family properties (Cervero 1997; Cervero and Duncan 2002a, 2002b), although the differences are not significant.

The table also gives some simple comparison tests of the means for each of the categories. The *t*-test statistic in the table is a group-wise mean equality test. In each category the equality test is done against the reference group in each category. The null and the alternative hypotheses of the test are given as follows:

$H_0 : Mean(ES \mid ref) - Mean(ES \mid j) = 0$, and

$H_a : Mean(ES \mid ref) - Mean(ES \mid j) \neq 0$.

where, ES is the effect size of the studies, j is an identifier of a group in the same category as the reference (ref). For instance, for the category 'type of railway station' light rail transit stations are the reference, and the other types of stations are compared with this. The test is performed under the assumption that population variance is unique. The t-test statistic is given by:

$$T = \frac{\overline{ES}_{ref} - \overline{ES}_j}{\sqrt{\frac{n_{ref} + n_j}{n_{ref}.n_j} \times \frac{(n_{ref} - 1).s_{ref}^{\ 2} + (n_j - 1).s_j^{\ 2}}{n_{ref} + n_j - 2}}} \tag{3}$$

Table 2.5: Descriptive summary of railway station proximity impact on property value (measured as a relative change)

	Effect within 1/4 mile						Effect per 250 metres					
	N	Min	Mean	Max	Stdev	t test	N	Min	Mean	Max	Stdev	t test
Overall	55	-0.619	0.081	1.452	0.263		57	-0.128	0.026	0.387	0.065	
Property Type												
Residential	42	-0.193	0.046	0.429	0.118		44	-0.038	0.019	0.134	0.035	
Commercial	13	-0.619	0.191	1.452	0.496	-1.773	13	-0.128	0.048	0.387	0.122	-1.428
Residential Properties												
Single Family[b]	29	-0.187	0.048	0.370	0.098		31	-0.031	0.024	0.134	0.036	
Condominium	6	-0.193	0.043	0.429	0.209	0.093	6	-0.038	0.008	0.084	0.041	0.963
Multi-Family	7	-0.086	0.040	0.291	0.121	0.196	7	-0.021	0.005	0.039	0.019	1.338
Type of railway stations												
LRT[b]	16	-0.072	0.071	0.302	0.093		18	-0.014	0.027	0.134	0.040	
HRT	20	-0.619	0.021	0.370	0.199	0.933	20	-0.128	0.009	0.099	0.043	1.292
CRT	15	-0.270	0.187	1.452	0.425	-1.093	15	-0.056	0.053	0.387	0.105	-0.977
BRT	4	-0.149	0.017	0.200	0.147	0.942	4	-0.031	0.003	0.042	0.030	1.104
Model												
Linear[b]	43	-0.619	0.079	1.452	0.291		45	-0.128	0.023	0.387	0.071	
Semi-Log	8	-0.187	0.085	0.370	0.157	-0.049	8	-0.006	0.037	0.099	0.040	-0.543
Log Linear	4	0.050	0.085	0.137	0.040	-0.037	4	0.016	0.034	0.046	0.014	-0.356
No Accessibility [b]	12	0.005	0.127	0.370	0.109		13	0.002	0.049	0.134	0.039	
Accessibility	43	-0.619	0.067	1.452	0.292	0.695	44	-0.128	0.019	0.387	0.070	1.485
No Demographic [b]	16	0.005	0.110	0.370	0.098		17	0.002	0.043	0.134	0.036	
Demographic	39	-0.619	0.069	1.452	0.307	0.526	40	-0.128	0.019	0.387	0.073	1.277
Time												
Up to 1990[b]	13	0.005	0.095	0.370	0.097		14	0.002	0.045	0.134	0.035	
After 1990	42	-0.619	0.076	1.452	0.297	0.226	43	-0.128	0.019	0.387	0.071	1.308

[b] = base group used as reference in the category. None of the equality tests are significant.

2.3.4 Random effect meta-regression model

Meta-analysis tries to explain variation in effect sizes by means of determinants as incorporated in Equation (2). In the literature, meta-regression is used in four different approaches: fixed effects; random effects; control rate; and Bayesian hierarchical modelling

(Morton et al. 2004). Fixed effect models assume that these estimates are random draws of one true value. The effect sizes included in the meta-analysis represent the estimates of the true value for the study with some degree of imprecision. Thus, the variance in the meta-analysis only comes from sampling error. However, substantial heterogeneity among the estimates can be an indication that the true effect value in the estimates is not unique. In such a situation Higgins and Thompson (2004) have indicated that fixed effect meta-regression models suffer from false positive results compared with the conventional regression model. The use of random effect models is believed to reduce spurious findings. In our case, the standard Q-statistics for the homogeneity test shows that the effect sizes of railway station proximity on property values show substantial heterogeneity[2]. This justifies the use of a random effects model for the meta-analytical procedure. The random effects model assumes that the variance associated with each effect size has two components: the within study variance and the between-studies variance.

In this chapter we apply the random effect meta-regression model to explain the variation in the effect sizes of the railway station proximity effect on property values. The variance of the effect size in this modelling approach is the sum of the two variance components: namely, the within-study variance (σ_i^2) and the between-studies variance (τ^2) components. Thus, the weight for each of the effect sizes is the reciprocal of this total variance ($w_i = 1/(\sigma_i^2 + \tau^2)$). The estimation procedure of the regression model proceeds in two stages. First, the between-studies variation measure (τ^2) is determined. Second, using the updated weight considering the within-study and between-studies variation, the regression analysis is performed. The regression equation estimated in this chapter is given in Equation 2. The Stata-based meta-regression routine (*metareg*) is used to run the estimation. An important feature of the random effect meta-regression is that R^2 is not reported; instead the τ^2 is reported. In Figures 2.2 and 2.3 the effect sizes used in our analysis are plotted against the corresponding standard errors of the effect sizes. Both graphs show a similar pattern, although the scale is different because of the different distance measures used.

---

[2] The homogeneity test's Q-statistic is given by $Q = (\sum w_i ES_i^{\ 2}) - (\sum w_i ES_i)^2 / \sum w_i$. w_i is the weight of the effect size (ES) of study i, given by the inverse of the variance. Q has a chi-square distribution. For the data in the analysis Q=1212, where the critical value for 5% and 56 degrees of freedom is 74.5. This indicates the effect sizes have substantial heterogeneity. This calls for a random effect model of estimation.

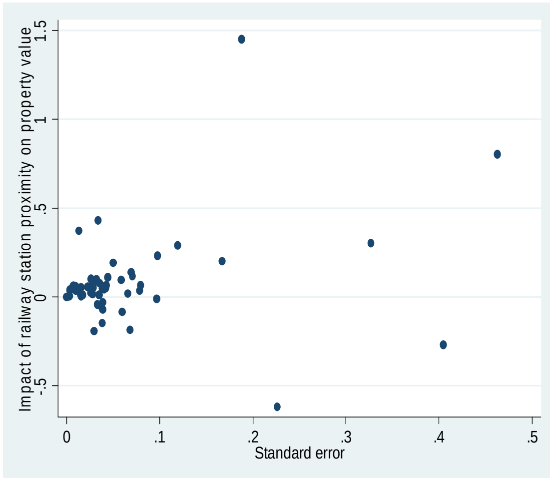


Figure 2.2: Plot of the railway station proximity effect for properties within ¼ mile of the station against the standard error of the estimates

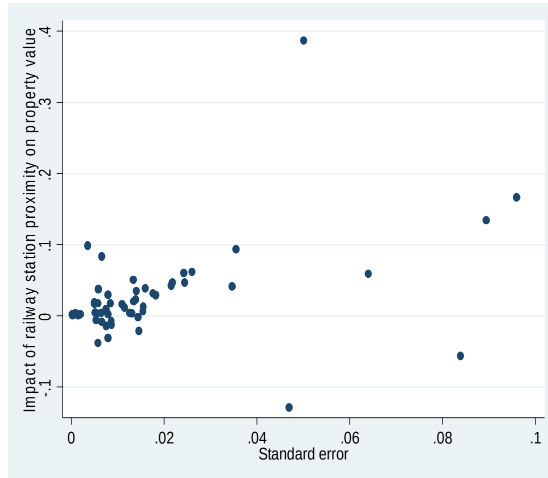


Figure 2.3: Plot of the railway station proximity effect for every 250m coming closer against the standard error of the estimates

2.3.5 Estimation Results

To explain the variation in the findings of the railway station proximity effect on property values by various study characteristics, we performed two estimations. As indicated in Section 2.3.2, the first estimation explains the impact of station proximity on the value (rent) of properties located within ¼ mile (402m) of the station. The impact is measured as the relative change in property values. The second estimation explains the impact of station proximity on property values (rents) for every 250 metres closer to the station. The explanatory variables for the two estimations are given in Table 2.4 above. The outputs of the random effect meta-regression model based on 55 effect sizes are given below.

1. Local effect of railway proximity:

In this case, the dependent variable is the effect of railway station proximity on properties located within ¼ mile distance of the station, compared with properties located outside this range. This measures the most localized impact of railway station accessibility on property values. The distance category is common to many studies. In addition, this range represents locations within walking distance. The random effect estimation results for this specification are given in Table 2.6 below.

Table 2.6: Effect of railway on property values within ¼ mile compared with other locations beyond

variable	coefficient	standard error	z-value	p-value
constant	0.087	0.071	1.240	0.215
commercial property	0.122	0.063	1.950	0.051*
heavy rail transit (hrt)	0.009	0.051	0.180	0.857
commuter rail transit (crt)	0.141	0.063	2.260	0.024**
bus rapid transit (brt)	-0.015	0.080	-0.180	0.856
semi-log specification (semilog)	-0.005	0.070	-0.080	0.940
Log-linear specification (loglinear)	-0.005	0.095	-0.050	0.956
accessibility variables	-0.187	0.094	-2.000	0.046**
demographic variables	0.055	0.091	0.600	0.545
time of data after 1990	0.029	0.061	0.480	0.633

No. of studies = 55.

τ^2 estimate = 0.0153.

* = significant at the 10% level; ** = significant at the 5% level; *** = significant at the 1% level

In Table 2.6 above, we see the τ^2 is greater than 0 (which would be the outcome if the fixed effect assumption held). This shows that there is substantial variation between the effect sizes (ES) of the studies. This confirms the justification for the use of a random effect model. Railway station proximity has a higher effect on commercial property compared with residential properties. The gap between the price within the ¼ mile zone and the remaining part of the city is larger for commercial property than it is for residential property. To be more precise: it is 12% larger. Table 2.5 shows, that while the price gap between the railway station zone and the rest is about 4.2% for the average residence, it is about 16.4% for the average commercial property.

The coefficients for heavy and commuter rail transport are positive, indicating that the effects of heavy and commuter rail transport on property values are greater than those of light rail transport (the base line in the estimation). Heavy railway transit stations have a 0.9% higher effect on property value compared with the effect of light rail transit stations. However, the significance level for this variable is low. On the other hand, a commuter rail transit station has a significantly higher effect on property values compared with light rail transit stations. It has an effect as big as 14.1% higher than the effect of light rail transit stations. This finding is consistent with the a priori expectation, and reflects the fact that commuter railways usually have wider service coverage (i.e. a larger catchment area).

The inclusion of other accessibility factors (highway, freeway) in the underlying studies significantly reduces the level of the reported station impact on property values (the reference group is the "no alternative accessibility variable in underlying study"). This shows that highways and freeways are also important determinants of property values (rents), next to railway station proximity. When both railway and highway accessibilities are included in the models (railway station and other modes), the effect on property values is 'shared' between the two different modes. Models with highway accessibility on average report 18.7% lower railway station proximity effects on property value than models excluding highway accessibility. The type of model specifications, temporal features, and demographic characteristics in the underlying studies show no significant explanatory power for the variation in the effect sizes of the studies.

2. Global effect railway station distance

In addition to the localized effect measure discussed above, effect sizes of railway station proximity for a wider range of distance from the stations were determined. Distance is now represented as a continuous measure. The effect sizes used in the estimation here represent the effect on property values of coming 250 metres closer to the railway station. There is no special reason for the choice of the 250m measure. The dependent variable values are given in percentage units. We use the term 'global effect' since the linear effect measure accounts for the whole range of distances to the railway station covered by the studies. The estimation results are given in Table 2.7 below.

The estimation shows a significantly negative coefficient for commercial properties as compared to residential propertied. This means, keeping other things constant, for every 250 m close to a station one comes the effect of the station on commercial properties is 2.3% lesser than on residential properties. To put this in perspective, if the value of residential properties increases by 2.4% for every 250 m closer to a railway station, the value on commercial properties increases by only 0.1%.

Table 2.7: Impact of railway station proximity on property values for every 250m closer to the station

variable	coefficient	standard error	z-value	p-value
Constant	0.049	0.004	11.870	0.000***
commercial property	-0.023	0.005	-4.310	0.000***
heavy rail transit (hrt)	0.000	0.001	-0.590	0.557
commuter rail transit (crt)	0.030	0.004	7.380	0.000***
bus rapid transit (brt)	-0.010	0.005	-2.150	0.032**
semi-log specification (semilog)	0.014	0.004	3.890	0.000***
log linear specification (loglinear)	0.002	0.009	0.260	0.796
accessibility variables	-0.014	0.006	-2.510	0.012**
demographic variables	-0.025	0.007	-3.280	0.001***
time of data after 1990	-0.008	0.005	-1.590	0.112

No of studies = 57.

τ^2 estimate = 1.1e-07.

* = significant at the 10% level; ** = significant at the 5% level; *** = significant at the 1% level

The results from this estimation are in some respects different compared with the localized effect analysis discussed above. This shows that different spatial considerations in addressing railway station proximity have a different impact implication for some study characteristics. We see a change in the sign for the effect on commercial properties compared with the residential properties. This means that the rent curve as a function of distance to the railway station is steeper for residential property than for commercial property. This is a remarkable result since the opposite was found for the local effects of stations (see Table 2.7). The reason for this difference is that, for commercial property, the direct proximity effect dominates: only when the office is within walking distance of the station (about ¼ mile) does it benefit, otherwise the station is of little use, and hence the rent curve is rather flat. The flat nature of the rent curve for distances further away than ¼ mile apparently dominates the pattern here. Since dwellings are located at the trip-origin side of stations, the car may also be used as an access mode and this gives the rent curve a higher slope across the whole range of distances.

Bus rapid transit stations (BRT) also have a significantly lower effect on property values than light railway stations. The signs of the effects for commuter rail transit and the inclusion of the accessibility variable are not affected. Commuter railway stations have on average a 3% higher effect on property values for every 250m closer to the station as compared with the effect of light railway stations. In addition to the presence of the accessibility variable, the presence of demographic variables in the studies also lowers the reported railway station effect on property values. This again underlines the importance of omitted variables bias in this type of studies.

2.4. CONCLUSION

The impact of railway station proximity on property values has received wide attention in the economics literature. Several empirical studies have tried to quantify this effect. However, the conclusions are not uniform. The aim of this chapter was to find a systematic explanation for the variation in the findings on railway station impact. We established that the different features of the study settings could explain these variations. We have tried to relate the variation to six categories of variables. These are: type of property under consideration; type of railway station; type of model used to derive the valuation; the presence of specific variables related to accessibility; demographic features; and the time of the data. The impact of railway stations on property values differs across property types. Generally speaking, railway stations are expected to have a higher positive effect on commercial properties compared with residential properties for relatively short distances from the stations. Among the four types of railway stations, commuter railway stations are expected to have higher impacts on the property values. The presence of accessibility and house quality variables is expected to have a negative effect on the magnitude of the impact of the station on the property values reported. We do not have a prior expectation of the impact of a specific functional form on the effect size for station proximity. This chapter presents two estimations based on two proximity considerations. First, we consider a local station effect by analysing the effect of a railway station on properties within a range of ¼ mile from the station. Second a more global effect is analysed based on a continuous measure of distance for a wider distance range.

Throughout the analysis, commuter railway stations show a significantly higher impact on property values compared with light or heavy railway/Metro stations. Their higher service coverage adds to the attraction of the area surrounding the stations. In addition, the number of commuter railways station is (relatively) low compared with light and heavy railway/Metro stations. The effect of a railway station on different property types is subject to spatial considerations. The effect on commercial properties is generally local. On average, commercial properties within ¼ mile of the station sell or rent at 12.2% higher than residential properties in the same distance range. Whereas the price gap between the railway station zone and the beyond is about 4.2% for the average residence, it is about 16.4% for the average commercial property. Note that the reference group for both properties is the set of properties that are situated beyond the ¼ mile range from the railway station. However, when considering global effects, the relative impact is reversed. On average, for every 250m closer to the station, the effect of the railway station is 2.3% higher for residential properties compared with commercial properties.

A given area can be made accessible by a number of modes (railways, car, etc.). Each mode will improve the accessibility of the region independently. All of the studies used in the meta-analysis analyse the (isolated) effect of a railway station on property values. When other accessibility modes are included in the underlying studies, railway stations generally have a lower impact on property values. Although both highways (freeways) and stations may increase property values, there is a negative correlation between the two effects; when one of these is present in a study, the effect of the other is diminished. Thus, we find an example of omitted variable bias: when highway accessibility is not explicitly addressed, railway impacts on property values tend to be overestimated especially in the continuous space specification.

The findings of this chapter highlight the difference in the railway-station effect between residential and commercial properties; the varying degrees of impact exerted by different types of railway station; and the importance of other transport modes in determining property values, together with railway accessibility. All these issues will be taken into consideration in this thesis. In the following chapter (Chapter 3), we give an urban model with two transportation modes. In order to better distinguish railway stations from each other, measures of quality and general railway accessibility are introduced. These will be the subject matters of Chapter 5 and Chapter 6, respectively. Estimations will be given for both residential house prices and commercial rent levels in the context of Dutch real estate markets (see Chapters 4, 7 and 8).

APPENDIX 2AI: Deducing the continuous railway station effect from discrete measures

The basic methodology for this was to linearize the impact over the different segments. For this method to work, it is required that the studies used at least three segments, including the reference segment in their analysis. Based on the assumptions described in Section 5.2.2, we can fairly say that the impact of railway station proximity on properties at the average distance of the segment from the station represents the effect of the station on the segment. The average distance of each segment is given by d = a + 2/3*(b-a), where "a" is the distance of the inner circle to the station, and "b" is the distance of the outer circle of the segment to the station. The reference segment's (the segment with value 100) outer circle is specified based on assumption 1 unless otherwise specified in the underlying studies. This gives us two corresponding variables (distance and value) for which we can estimate the percentage change in property values per unit of the distance measure using a semi-log specification:

$$\ln(\textit{value}) = a_0 + b_1 \times D,$$

where, *value* is the value of properties at distance D from the railway station. The value of the coefficient b_1 measures the percentage change in property values for a unit change of distance

Chapter 3

3 The effects of railway investments in a polycentric city: a comparison of competitive and segmented land markets[3]

3.1 INTRODUCTION

The car has gradually become the dominant transport mode in most cities in developed countries. However, there still are cities, such as London, Paris and New York, where a large part of the workers use public transport. Therefore, to make a proper analysis of land and labour markets in such cities, both transport modes should be considered. Many cities started with a clear monocentric structure. During the course of time, however, a gradual de-concentration process took place, leading to a decreasing dominance of the original centre. But, in some cases, edge cities have developed, implying the emergence of additional centres of commercial activity in a metropolitan area (Garreau 1988). In other cases, the gradual growth of small and medium-sized cities led to the evolution of large metropolitan areas consisting of overlapping urban areas that were formerly independent. In both these cases of city evolution, the original monocentric urban model no longer applies. This chapter sets out to study both phenomena in an urban model which deals with the combination of multiple transport modes and multiple centres of economic activity.

In relation to the type of centre and mode of transport assumptions, we can logically distinguish four categories of urban models: monocentric-unimodal transport; monocentric-multimodal transport; polycentric-unimodal transport; and polycentric-multimodal transport models. Most of the literature is in line with the first category, although recently, more and more studies that use the polycentric city setting have appeared. But they are still dominated by the unimodal transport assumption. This chapter is a study in the last of the four above-mentioned categories. It deals with an urban model of two centres and two transport modes. The goal of this chapter is to develop an urban model with an emphasis on the impact of investment in transportation on the real estate market. Thus, this chapter assesses the effect of

---

[3] This chapter is based on Debrezion, Ghebreegziabiher, Eric Pels and Piet Rietveld (2006). “The Effects of Railway Investments in a Polycentric City: a comparison of competitive and segmented land markets”. *Environment Planning A*, Forthcoming.

investment in rail on the spatial rent structure. Three levels of railway investments: namely, no rail investment; partial railway investment where only one of the two centres is connected by railway; and complete railway investment in which both centres are connected by railway, are compared. Moreover, this helps to assess the effects of an additional transport system (in this case: rail) on urban growth in general, and the growth of particular sub-centres. In addition, this chapter assesses the effect of partial railway investment on the competitive positions of centres within the city. Finally, the chapter addresses an institutional issue, i.e. the extent to which a regulated land market would lead to different results. In particular, we address the following question: Which institutional setting (competitive versus segmented market) leads to the highest rent increases as a result of investments in rail infrastructure? For each of the three levels of railway transport investment, we consider three situations concerning the land market regimes in the centres: a competitive land market situation in both centres; a segmented land market situation imposed in both centres; and a mixed land market in which one centre has a competitive land market while segmentation is imposed in the other.

In subsequent sections, we briefly discuss the relevant literature (Section 3.2), the specification of our model (Section 3.3), and the equilibrium conditions of the model (Section 3.4). We introduce the model for alternative land markets in Section 3.5. The model simulation and results are discussed in Section 3.6. Finally, we conclude and suggest directions for future research (Section 3.7).

3.2 LITERATURE REVIEW

The relation between land values and proximity of land to employment centres has been addressed extensively in the literature. The monocentric circular city has received most attention. However, in many parts of the world, especially in Western Europe, Japan and the U.S., metropolitan areas are increasingly assuming polycentric structures. The Randstad in Holland, the Rhine-Ruhr metropolitan area in Germany, the Flemish Diamond in Belgium, the Glasgow-Edinburgh corridor in the UK, the Padua-Treviso-Venice region in Italy, Southern California in the U.S. and the Kansai area in Japan are probably the most frequently mentioned polycentric structures (van der Wusten and Faludi 1992; Dieleman and Faludi 1998; Batten 1995; Musterd and van Zelm 2001). But, even though polycentric urbanization started earlier in Europe than in the U.S., its pace was slowed in Europe as a result of conservative urban policies.

A common starting point in the literature is that transportation is the prime factor in shaping the urban structure (Clark 1958; Clark and Kuijpers-Linde 1994). Besides changes on the transportation side, changes on the production side (agglomeration and productivity effects) are responsible for determining location patterns and thus shaping the urban economy (Glaeser and Kahn 2004). Fujita et al. (1999) theoretically explained the effects of agglomeration on the optimal location of firms in relation to the location of a historic centre. In a linear city of unit length, the optimal location of a new plant will be in the historic centre for a wide range of cases. Nevertheless, at times, the optimal location of the plant can be different from the historic centre. The trade-off between agglomeration effects and transportation costs explains the coexistence of multiple centres in a city. Modarres (2003) found, for Los Angeles County, that sub-centres contain one-third of the county's employment. However, the public transport network structure appears to serve these sub-centres inadequately. This shows that, in this case, the formation of a polycentric urban structure was not in response to the development of public transport in the first place. However, the increasing use of private cars is believed to be instrumental in shaping the present dispersed urban structure (Clark 1958; Glaeser and Kahn 2004). In addition to the use of cars, Sivitanidou (1997) showed that the recent information revolution is also contributing towards the weakening of spatial links between commercial activities and large business locations, thereby leading to increasingly dispersed business locations.

Even though polycentricity simply implies the presence of multiple centres in an urban area, there is no proper identification procedure (Anas et al. 1998). For practical purposes, areas can be treated as centres in terms of variables such as employment density, population density, property values and travel patterns. Several authors have tried to propose ways of identifying centres in cities by both parametric and non-parametric methods. However, these still remain essentially subjective. The main methods used to identify sub-centres are: the residual method of McDonald (1989); the employment density cutoffs method of Giuliano and Small (1991); and the employment smoothing estimation procedure of Craig and Ng (2001). Later, McMillen (2001) developed a two-stage centre identification procedure, which incorporates concepts of McDonald (1989), and Craig and Ng (2001). In the first stage, candidate centres are identified through the analysis of the residuals of a smoothed employment density function. The second stage assesses the significance of the identified centres in influencing local employment densities. This reflects the definition that centres are sites which result in a significant rise in employment densities after controlling for the historic centre (the CBD). Apart from calculating an employment density indicator, Musterd and van Zelm (2001) discussed various ways to define a polycentric city structure. Both spatial structure and the existence of intricate network-type interactions should be present before considering an urban area to be a polycentric unit.

Several studies have addressed the effect of urban spatial structure on property values. This will also be the main focus of this chapter, which attempts to answer the question: How does the polycentricity of an urban area shape the land rent structure? The value of a centre is capitalized in the form of land rents. In addition to the predetermined centre in the urban models, other studies, without explicitly referring to the centre(s), have concluded that the rent gradient peaks around the most valuable location in the urban spatial structure. Indirectly, these peaks are also used to identify the centre(s) of the city. However, in this sense, the monocentric assumption is in reality a very simplistic assumption. Therefore, over the years, attempts have been made to develop urban models in the context of polycentric situations (pre-specified and non-pre-specified locations).

A comprehensive general equilibrium polycentric urban model was developed by Anas and Kim (1996). Without scale economies of shopping, production is dispersed in the city with rent, wage, and commodity price and density gradients peaking in the centre of the space. One of the models on property values in a bicentric city was developed by Sivitanidou and Wheaton (1992). Special attention was given to the centres' production cost difference and commercial land market regulation. The main finding of the chapter was that cost advantages are capitalized in commercial land rent and wages (and wages, in turn, capitalized in residential land rents). The level of capitalization of production cost advantages in commercial land rent becomes higher in the regulated commercial land market compared with the competitive market. In this chapter, we extend Sivitanidou and Wheaton's (1992) model by introducing an additional mode of transport (rail: fast mode with discrete access points) running through the bicentric linear open city. In the model, households and firms interact via the exchange of labour and wages. The differences between our model and that of Sivitanidou and Wheaton (1992) concern aspects such as: the introduction of a second transport mode;

endogenous land consumption by households; endogenous density of settlements; and endogenous wages for the two centres.

3.3 THE STRUCTURE OF THE MODEL

3.3.1 Bicentric-Bimodal urban structure: model description

In this section, we introduce an equilibrium urban land use model. The city in our model has two centres that both function as employment and production centres. Labour is employed from households living in the residential areas of the city. Homogeneous households arrive at either of the two centres and supply labour. The inputs in the production process constitute labour and commercial floor space. In the production process of the centres, we assume a fixed ratio between labour and floor space. Floor space is prepared in a cost-minimizing fashion from land rented at the commercial land rent rate and capital rented at some market rent of capital. The output follows a fixed proportion, constant-returns-to-scale technology and is exported at a given price in a fully competitive market.

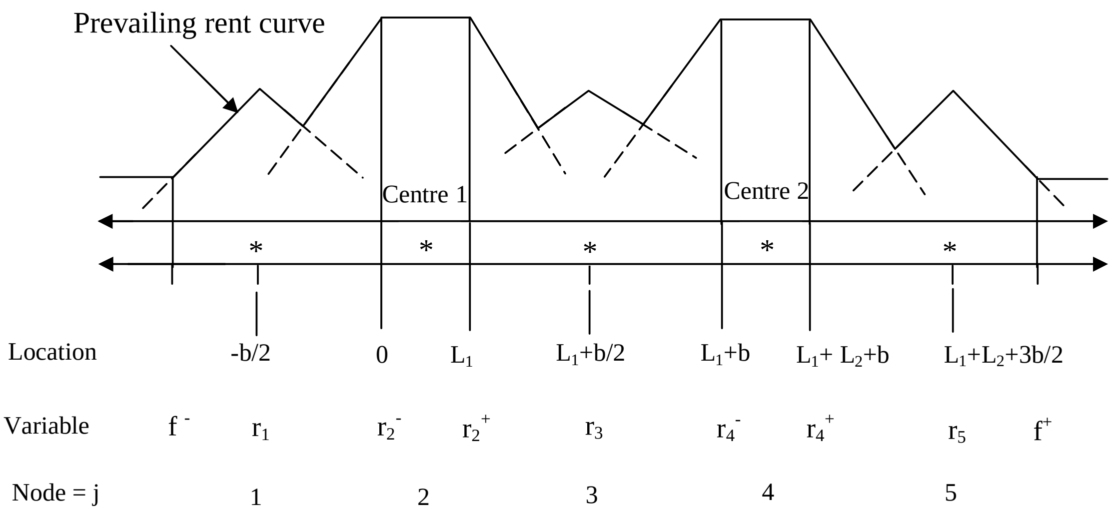
The households are assumed to have a well-behaved utility function with residential land and non-land consumption goods as its components. By travelling to one of the employment centres, households acquire an endogenously given wage. No other income sources are considered. The residential land rent has a bid nature. The price of non-land consumption goods is taken as a numéraire (unity). All commercial and residential rents are absorbed by absentee landowners. We further assume that the city we deal with is open: households can freely migrate into or out of the city. The households enjoy the national utility level u which is bounded from below by the supreme utility level (a level of utility that guarantees the existence of the city) (Fujita 1989). Thus, all households in the city enjoy a given utility level that is equivalent to the level of utility enjoyed by the households outside the city in the economy.

Two modes of transport operate in the open city: a “slow” mode (car) that is accessible from anywhere in the city, and a “fast” mode (train) accessible from certain fixed stations. The distinction between the “slow” and “fast” modes does not only relate to the time cost of transportation, but rather to the generalized transportation cost structure. The fast mode results in some sort of cost saving, and is thus termed “fast”. Thus, the cost per distance unit of transportation for the fast mode (train) is lower than that of the slow mode. In our linear city model, as well as at the centres, we assume three additional stations at a distance of *b/*2 miles from the edges of the centres (see Figure 3.1 below). We assume transportation costs inside the centres are zero. The households use the slow or fast mode or a combination of both in a cost-minimizing fashion to reach the centre, where they earn net income (wages minus transport costs). There are three possible situations: a resident at the location of a railway station uses only the railway mode; households at the location near a railway station use a car to the station and transfer to the railway for commuting to the employment centre; and workers at locations near employment centres use only a car for commuting.

The exogenous parameters of the model are the following. On the consumer side, we have the national utility level, the price of non-land consumption goods, and transportation costs of the two modes. On the producer side, we have the floor space requirement per worker, and the cost of capital. Lastly, for the spatial structure of the city we have the width of the city and the distance between the nodes. The values of the exogenous variables used in the simulation are given in Table 3.3. Analysis is provided for two regulatory alternatives for land market situations: competitive and segmented markets.

3.3.2 Notation and definition of variables

The general layout of the city structure is depicted in Figure 3.1. The two centres occupy a significant amount of land for commercial purposes in the urban area. The two centres are *b* distance units away from the edges. The left edge of Centre 1 is taken as the origin of the linear city. In Figure 3.1, the second row gives the variable representing the location of some critical points in the linear city such as the fringes of the city (f^- and f^+), the edges of the centres (r_2^- and r_2^+ for Centre 1 and $r4^-$ and r_4^+ for Centre 2) and the location of railway stations (r_1, r_3 and r_5). L_1 and L_2 represent the size of the two centres, respectively. According to the land market assumptions, they are exogenous to the model or are endogenously determined in the model. This is dealt with in detail in the next section. Table 3.1 introduces the variables and notation used in the next section to formulate the model. These variables are discussed in more detail below when the model details are explained.



* The location of the railway stations in the linear city

Figure 3.1: Layout of linear city with two commercial centres, each with its own railway station and three additional railway stations

Table 3.1: Model variables

Variable	**Description**
I	The set of employment and production centres $I = \{1, 2\}$
J	The set of transport nodes $J = \{1,2,..5\}$, where $I \subseteq J$
r	A variable location in one dimensional space
r_j	Location of node j, $j \in J$ in one dimensional continuous space
u	Utility level
w_i	Wage at Centre i, $i \in I$
Y_j^*	Artificial income at node j, $j \in J$
R_A	Rent for agricultural land (the basic land rent)
RC_i	Commercial land rent at Centre i, $i \in I$
r_c	Rent of capital ($/sq. ft.)
RF_i	Rent for floor space at Centre i, $i \in I$
$\Psi_j(r,u)$	A function of residential bid-rent land corresponding to node $j \in J$ at location r and u level of utility
$RR(r,u)$	A function of prevailing residential rent per unit lot size of land at location r for utility level u.
S(r,u)	A function of size of residential land consumed by household at location r, the max bid lot size

Z(r,u)	A function of non-land goods consumed at location r. (taken as numéraire with unit price).
$U(z,s)$	Utility as a function of z level of non-land goods and s level of land consumed
k_s	Transportation cost per unit distance for the slow mode
k_f	Transportation cost per unit distance for the fast mode
$T(r)$	Transportation cost function from location r to the destination centre (node)
N_i	Number of employees at centre i, $i \in I$
Qs_i	Floor space at Centre i, $i \in I$
L_i	Area of commercial land at centre i.
K_i	Amount of capital employed at Centre i, $i \in I$
a_s	Floor space per workers ratio (Qs_i/N_i), $i \in I$
CE_i	Other exogenous costs per worker in Centre i
CA	Production cost advantage for Centre 1 (=CE_2-CE_1)
prt	Productivity per worker (units/worker).
$l(r,u)$	Distribution of land in the city. In the linear constant unit the width of the city is given by $l(r,u)$ = 1.
ρ(r,u)	Density of households of a city at location r, corresponding to level of utility u.

3.4 MODEL DETAILS

3.4.1 Household behaviour: utility maximization

Besides the agricultural land use that starts at the outermost fringes of the city, urban land is devoted to residential and commercial use. The assumption of a land market without any imperfections guarantees that commercial land rent always outbids residential land rent in the employment centres. For residential land use, the trade-off between transport costs and other consumption opportunities leads to a downward-sloping bid-rent curve from the edge of the centres. Thus, the land-rent curve is an envelope of the curves, as depicted in Figure 3.1 above.

We start with the derivation of the residential bid-rent function. The bid-rent is defined as the maximum rent per unit of land that a household, at a location *r* that travels to a specified employment centre to get an income *Y*, can pay while achieving a certain utility level u. The bid-rent function in the city therefore is a function of the distance and the utility level enjoyed by the households (Fujita 1989), which can be written as:

$$\Psi(r,u) = \max_{z,s} \left\{ \frac{Y - T(r) - z}{s} \mid U(z,s) = u \right\} , \tag{1}$$

where, $\Psi(r,u)$ is the residential land bid-rent function, for a household at location *r* enjoying a given exogenous level of utility u. $U(z,s)$ is the utility function, where *z* is the composite consumption good of the household that has a unit price, and *s* is the land lot size per household. The household incurs transportation costs *T(r)* which is a function of the location *r* in reference to the location of the employment centre, and receives a level of income *Y*.

Equation (1) can be rewritten by expressing the amount of composite consumption goods of the household in terms of lot size of land and utility:

$$\Psi(r,u) = \max_{s} \left\{ \frac{Y - T(r) - Z(s,u)}{s} \right\} . \tag{2}$$

For a fixed utility level u, the first-order condition for maximizing the right-hand side of Equation (2) occurs at the point where the marginal change of the function with respect to *s* is zero. This leads to the relation:

$$-\frac{\partial Z(s,u)}{\partial s} = \frac{Y - T(r) - Z(s,u)}{s}. \tag{3}$$

At the optimal choice of s the right-hand side of (3) equals the bid-rent $\Psi(r,u)$:

$$-\frac{\partial Z(s,u)}{\partial s} = \Psi(r,u). \tag{4}$$

This means that a marginal decrease in the consumption of non-land composite consumption goods due to an increase in the consumption of lot size of land is equal to the bid-rent of land.

For simplicity and ease of derivation, we assume that the utility function of the household has the following functional form:

$$U(z,s) = \alpha \ln z + \beta \ln s; \quad \alpha + \beta = 1. \tag{5}$$

It can be shown that $Z(s,u) = s^{-\beta/\alpha} e^{u/\alpha}$. Solving the maximization problem in (2) using the condition in (3), the following residential bid-rent function can then be derived (Fujita, 1989, p. 322):

$$\Psi(r,u) = \alpha^{\alpha/\beta} \beta (Y - T(r))^{1/\beta} e^{-u/\beta}; \tag{6}$$

Given the bid-rent level for the price of land, the lot size level that optimizes utility is given as:

$$S(r,u) = \beta (Y - T(r)) / \Psi(r,u) = \alpha^{-\alpha/\beta} (Y - T(r))^{-\alpha/\beta} e^{u/\beta}. \tag{7}$$

The density of settlement (ρ) is given by the inverse of the max-bid lot size, and gives the number of households per unit lot size area:

$$\rho(r,u) = 1/S(r,u). \tag{8}$$

(a) Income at nodes

The above bid-rent function can easily be applied in the case of a monocentric unimodal city. However, in the present case with two modes and two commercial centres, some further steps are needed. First, the household's transport cost curve $T(r)$ has to be derived from the modal choice model. Second, the income level Y is no longer unique since wages may be different in the two nodes. As will be explained below, in order to keep the model tractable, we introduce the transport costs related to the fast mode in the income variables.

Households travel to the employment centre that maximizes their net income. Because of the cost saving nature of the fast mode (rail), we observe three possible ways of commuting to the employment centres. First, households residing at the stations will directly use the fast mode to commute to the employment centre which maximizes their net income. Second, households residing around the stations will use the slow mode (car) to reach the stations and then take the fast mode to the employment centre which maximizes their net income. Third, households residing around the employment centres will directly use the slow mode to commute to the employment centre. Thus, the slow mode has two types of destinations: a transfer station or a real employment centre.

We now turn to the income levels earned in the various nodes. For people working in the two commercial centres, j= 2 and j= 4, and travelling by car to these centres the income equals the pertaining wage levels w_1 and w_2.[4] In order to relate the bid-rent analysis to our multi-nodal model, we introduce a pseudo-income variable for the other workers. First, we consider stations 1, 3 and 5 as pseudo centres with zero production. A pseudo-income is then attached to these pseudo-centres. These are equal to the net income that households residing at these centres get by commuting to the employment centres that maximizes their net income using the fast mode. Thus, we extend this income definition over all nodes (railway stations and employment centres) as given by Equation 9. The introduction of this pseudo-income variable helps to keep the model simple by internalizing the transport costs related to the fast mode in the income variable. We use the term 'artificial income' for the pseudo income attached to each of the nodes (Y_j^*), and it is defined as follows:

$$Y_j^* = \max_i (w_i - T_f(r_{ji})) \qquad j \in J \ \ and \ \ i \in I \ , \tag{9}$$

where, $T_f(r_{ji})$ is the transportation costs from node j to centre i by the fast mode. In our analysis we adopt the linear transportation cost function $T_f(r_{ji}) = k_f \times \| r_i - r_j \|$. Thus, the

[4] A slightly more general formulation would allow one of the commercial centres not to materialize because productivity is too low compared with the other node. This can easily be taken on-board in the present model formulation, but we decided not to do this because it would lead to more complex model formulations without adding substantial insights.

artificial income at node j equals the wage in the commercial centre towards which it is oriented, minus the transport costs to get there by fast mode. The equality of the artificial income of the two centres in the city with the real wages offered in the corresponding centres ($Y_2^* = w_1$ and $Y_4^* = w_2$) represents the coexistence of the centres as both production and employment centres. However, if one of the artificial incomes is higher than the real wages offered at the corresponding centre, this implies that the centre ceases to be a production centre. This means that this centre is dominated by the other centre: it serves as a mere transfer node to the dominant centre.

(b) Residential land rent:

Given the income level attached to each of the nodes in the city, we can safely assume that each node faces a downward-sloping residential bid-rent curve. It is a function of the utility level enjoyed by the households and the distance to the node. Households travel by the slow mode (car) to one of the nodes to work or make a transfer to the fast mode (train), depending on the nature of the node. If the node is an employment centre, households use the slow mode to reach the centre directly. In our model the two employment centres are indexed by $j = 2$ and $j = 4$ in the set of nodes. If the node is a mere railway station, households use the slow mode to reach the railway station and continue their trip to the employment centre by train. In the set of nodes, the transfer stations are indexed by $j = \{1,\ 3\ \textit{and}\ \ 5\}$ (see Figure 3.1). Generally, there are two distances involved. The first of these is the distance from the transfer railway station to the employment centre which is internalized in determining the income corresponding to the nodes. The second distance relates to the distance between the location of the households' residence and the nodes. We assume that the transportation cost by the slow mode is proportional to distance. Substituting the artificial income level at the nodes and the transportation cost of reaching the nodes by the slow mode in Equations 7, 8, and 9 gives the residential bid-rent, max-bid lot size and settlement density functions corresponding to each of the nodes:

$$\Psi_j(r,u) = \alpha^{\alpha/\beta}\beta(Y_j^* - k_s\|r_j - r\|)^{1/\beta} e^{-u/\beta}; \quad j \in \{1,2,\ldots,5\}; \tag{10}$$

$$S_j(r,u) = \beta(Y_j^* - k_s\|r_j - r\|)/\Psi_j(r,u); \quad j \in \{1,2,\ldots,5\}; \tag{11}$$

$$\rho_j(r,u) = 1/S_j(r,u); \quad j \in \{1,2,\ldots,5\}. \tag{12}$$

The residential bid-rent, residential lot size, and settlement densities are only defined in the residential areas of the city. Due to the bidding nature of rent, the prevailing land rent in the residential areas of the city is the maximum of the rent curves corresponding to each of the transport nodes in the linear city.

$$RR(r,u) = \max_j(\Psi_j(r,u)); \quad j \in \{1,2,...,5\}; \tag{13}$$

$$S(r,u) = \min_j(S_j(r,u)); \quad j \in \{1,2,...,5\}; \tag{14}$$

$$\rho(r,u) = \max_j(\rho_j(r,u)); \quad j \in \{1,2,...5\}. \tag{15}$$

(c) Commercial land rent

In the production process of the firm(s) operating at the employment centre, land is one of the inputs in the production of floor space. Because transportation costs inside the centre are assumed to be zero[5], a uniform land rent for commercial use is obtained. This assumption is not unnatural. In most urban models the transportation cost inside the CBD is ignored. Moreover, the size of the CBD is usually rather small compared with the rest of the metropolitan area. A consequence of the assumption is that households take the edge of the centres as a reference of the location of the employment centres[6]. At the edge of the centres, the commercial land rent curve takes over. This situation is guaranteed both under competitive and segmented land market assumptions as will be explained in Section 3.5. In the competitive land market situation, the commercial land rent outbids the residential land bid-rent curve. At the edge of the centre, the commercial land rents are equal to the corresponding residential land rents. On the other hand, the segmented market situation guarantees that the commercial land rent takes over whether it outbids the residential land rent or not. The commercial land rent function is derived from the producer behaviour in Section 3.4.2 below.

---

[5] It would also be possible to consider transportation cost inside the centres. However, this would strongly complicate the formulation of the model since it would lead to wage levels that depend on the location of work within the cluster. Workers at the fringe of the commercial area, i.e. those who travel by car, would earn (slightly) less than workers working closer to its centre because the latter would need compensation for the extra transport costs. Along similar lines, the workers who travel by train and whose job is close to the central station would have a lower wage than the workers who have to walk a certain distance from the station to the workplace. This would lead to a wage gradient in the commercial area that is low at the fringes and in the centre, and that has peaks in-between. This would imply that the rent in the commercial area would not be constant, and this also means that the densities would not be constant: high densities would be expected at the centre and at the fringes, and in-between lower values would be expected. Although such refinements would be interesting to study, we feel they would substantially complicate the analysis without major benefits in terms of additional insights into the themes studied.

[6] In fact their bid curve within the commercial centre would be flat, given the fact that transport costs within the centre are zero.

(d) Prevailing land rent and land use

In the model we distinguish three types of land use: commercial; residential; and agricultural. The agricultural land rent is given exogenous to the model. The uniform commercial land rent outbids the downward-sloping residential rent curve which starts at the edge of the centre. Thus, the prevailing rent curve at any point in space is the maximum of the residential, commercial and agriculture land rents, which can be written as:

$$R(r) = \max_{i \in I, j \in J} (\Psi_j(r,u), RC_i, RA) . \tag{16}$$

3.4.2 Producer behaviour: cost minimization

On the production side, the model incorporates the assumption of Sivitanidou and Wheaton (1992), in which the two centres make products that utilize labour and floor space and that are sold outside the city. Floor space at the centres (Qs_i) is produced by making use of capital (K_i) and land (L_i) with a constant-returns-to-scale Cobb Douglas production function, as given below:

$$Qs_i = K^{\delta} L^{1-\delta} . \tag{17}$$

Centres are assumed to have constant floor space demand per worker ($a_s = Qs_i / N_i$). Thus, given the price of capital (r_c), rent for floor space is given by:

$$RF_i = \frac{RC_i L_i + r_c K_i}{a_s N_i} , \quad i \in I . \tag{18}$$

where $a_s N_i$ is the total demand for floor space and the numerator is the total expenditure in floor space. The cost-minimizing input bundle for floor space is given by:

$$K_i = a_s N_i \left(\frac{\delta RC_i}{(1-\delta) r_c}\right)^{(1-\delta)} , \quad i \in I ; \tag{19}$$

$$L_i = a_s N_i \left(\frac{(1-\delta) r_c}{\delta RC_i}\right)^{\delta} , \quad i \in I . \tag{20}$$

Therefore, the long-run cost-minimizing floor space rent is:

$$RF_i = r_c^{\delta} \delta^{-\delta} (1-\delta)^{(\delta-1)} RC_i^{(1-\delta)}, \;\; i \in I. \qquad (21)$$

The total costs of production at each of the centres are the sum of wages, rent of floor space, and some other exogenous production cost. The other exogenous production costs may include all costs different from the ones already discussed (labour, land and capital). Examples are locally-varying tax payments, differences in local facilities, differences in costs of transport to the outside market, or agglomeration advantages and costs associated with the use of local public services. The difference between the exogenous production costs of the two centres reflects the level of cost advantage that one centre has over the other. The situation where $CE_2 - CE_1 > 0$ indicates Centre 1 has a cost advantage over Centre 2. Centres operate at a zero-profit level. Productivity per worker (prt) is assumed to be the same in both centres. The output of both centres is exported to a national market at a unit price. Wages in both centres are determined endogenously by the model:

$$prt - CE_1 - w_1 - RF_1 \times a_s = 0\,; \qquad (22)$$

$$prt - CE_2 - w_2 - RF_2 \times a_s = 0\,. \qquad (23)$$

In other words, the production equilibrium condition between the centres thus states that the costs per worker at each of the centres should be equal:

$$w_1 + RF_1 a_s + CE_1 = w_2 + RF_2 a_s + CE_2. \qquad (24)$$

The main theme of the paper by Sivitanidou and Wheaton (1992) was the effect of relative cost advantages on the spatial rent structures. In order to focus on the effect of railway investment on the spatial rent structure, our discussion of the simulation assumes the exogenous cost component.

3.4.3 Allocation of households to employment centres

The number of households in the city (N) is an important element in the determination of the equilibrium conditions. It is assumed that the city does not continue beyond the fringes of the residential areas. The total number of households in the city (given by Equation 27) is calculated as the integral of household density between the city fringes. The left (f^-) and right (f^+) fringes of the city are locations where the residential rent corresponding to the nearest node and the agricultural rent intersect. Solving Equations 25 and 26 gives the location of the fringes:

$$\Psi_1(f^-, u) = RA\,; \tag{25}$$

$$\Psi_5(f^+, u) = RA\,; \tag{26}$$

$$N = \int_{f^-}^{f^+} \rho(r,u)\,dr\,. \tag{27}$$

In order to determine the number of households commuting to each of the centres, we also need to know the location at which households are indifferent between commuting to both centres. The middle node (which lies halfway between the two centres) plays an important role in determining the position of the indifference location. Households arriving at this node will commute by the fast mode to the centre whose wage rate was used to determine the artificial income corresponding of this node (Y_3^*) (see Equation 9). If this wage corresponds to Centre 1, the indifference location is to the right of the middle node where the bid-rent curve corresponding to the middle node equals to the bid-rent curve corresponding to Centre 2. On the other hand, if the wage corresponds to Centre 2 the indifference location is to the left of the middle node at a point where the bid-rent curve corresponding to the middle node crosses the bid-rent curve corresponding to Centre 1[7]. Thus, the indifference location (*f*) can be given by:

$$\begin{aligned} &\Psi_3(f,u) = \Psi_4(f,u) \quad if \quad w_1 > w_2\,; \\ &f = r_3 \qquad\qquad\qquad\; if \quad w_1 = w_2\,; \\ &\Psi_3(f,u) = \Psi_2(f,u) \quad if \quad w_1 < w_2\,. \end{aligned} \tag{28}$$

Given that (f) is the indifference location between the centres, the number of households working in each centre is given by:

$$N_1 = \int_{f^-}^{f} \rho(r,u)\,dr\,; \tag{29}$$

$$N_2 = \int_{f}^{f^+} \rho(r,u)\,dr\,. \tag{30}$$

[7] At times it can also happen that the wages of the two centres are the same. Households arriving at the middle node can travel to any of the two centres. This opens a possibility for cross-commuting. In such a situation, the expected number of households commuting to the centres will be distributed equally, leaving the position of the node to be the effective indifference location.

3.5 ALTERNATIVE LAND MARKETS

In this section, we analyse the implications of the behaviour of the households and firms as described in Section 3.4 for two different institutional settings for the land market. We start with the usual assumption of a competitive land market, followed by the case of a segmented land market where the government intervenes by imposing constraints on the size of the commercial areas.

3.5.1 Competitive land market

In this model we assume households and firms freely bid against each other for land. At the edges of the centres, the commercial and residential rents are equal. The competitive land bidding ensures that landlords will eventually extract the maximum saving that the consumers may enjoy, given the utility level. Because, by assumption, there are no transportation costs inside centres, the commercial rent curve assumes a uniform pattern. Centres situated at pre-specified locations make no profit from production processes. The equilibrium conditions for the competitive land market are presented below in Table 3.2. Note that the system of 13 equilibrium conditions has 13 unknowns { RF_1, RF_2, w_1, w_2, RC_1, RC_2, f, f^-, f^+, N_1, N_2, L_1, L_2 }. The numerical results for this model are presented in Section 3.6.

Table 3.2: Equilibrium conditions

Description	***Equation***	***Previous reference***
Equality of residential bid-rent and commercial rent at the edges of employment centres	$\Psi_2(r_{2^-},u)=\Psi_2(r_{2^+},u)=RC_1$ $\Psi_4(r_{4^-},u)=\Psi_4(r_{4^+},u)=RC_2$	-
Cost-minimizing floor space function at the centres	$RF_1=r_c{}^{\beta}\beta^{-\beta}(1-\beta)^{(\beta-1)}RC_1{}^{(1-\beta)}$ $RF_2=r_c{}^{\beta}\beta^{-\beta}(1-\beta)^{(\beta-1)}RC_2{}^{(1-\beta)}$	(21), for $i\in I$
Zero-profit condition for production centres	$prt\times p-CE_1-w_1-RF_1\times a_s=0$ $prt\times p-CE_2-w_2-RF_2\times a_s=0$	(22) (23)
Commercial land for centres	$L_1=a_sN_1(\frac{(1-\delta)r_c}{\delta RC_1})^{\delta}$ $L_2=a_sN_2(\frac{(1-\delta)r_c}{\delta RC_2})^{\delta}$	(20), for $i\in I$
Number of households attending the centres	$N_1=\int_{f^-}^{f}\rho(r,u)\,dr$ $N_2=\int_{f}^{f^+}\rho(r,u)\,dr$	(29) (30)

Continued from Table 3.2

Left and right fringes of the linear city	$\Psi_1(f^-,u) = RA$ $\Psi_5(f^+,u) = RA$	(25) (26)
Indifference location between the two centres	$\Psi_3(f,u) = \Psi_4(f,u)$ if $w_1 > w_2$; $f = r_3$ if $w_1 = w_2$; $\Psi_3(f,u) = \Psi_2(f,u)$ if $w_1 < w_2$.	(28)

3.5.2 A segmented land market

In this land market situation, we impose a binding restriction on the commercial land area for one or both of the centres, such that $L_i = \text{rl} < L_i^*$, where L_i^* is the land area occupied by Centre i if no restriction is imposed on it, and rl is a fixed amount of land reserved for commercial land use. As a result, the commercial land prices are higher than would be possible under the competitive land market situation, because of the imposed scarcity. The restriction affects the commercial land rent, and the relative cost of land versus capital. Thus, at a centre with the commercial area restriction, the land rent is no longer determined by competitive bidding from residential land rent, but is instead based on the supply of commercial land rent. With a restricted supply of land, the commercial land rent increases, thus increasing the relative cost of land to capital in the centre. Leaving out the first two equations which are specifically related to the competitive market case from Table 3.2 above, the remaining 11 equations determine the equilibrium conditions for this model. The equilibrium condition in this market situation has 11 equations in 11 unknowns. The commercial land areas in the two centres, L_1 and L_2, are exogenous in this model.

3.6 MODEL SIMULATIONS, OUTPUT AND DISCUSSION

On the basis of the model above, we now present the simulation results for three transport mode situations and three land market regimes. With respect to the transport mode, we have looked at the unimodal case and, two bimodal cases: namely, partial and full bimodal. In the partial bimodal case, only Centre 1 is served by the fast mode from two stations b/2 miles away from its edges, in addition to the slow mode. On the other hand in the full bimodal case, both centres are served by the fast mode from the pre-specified stations, in addition to the slow mode. In each of the three levels of railway transport investments we can have three situations concerning the land market regimes in the centres. First, we can have competitive land market conditions in both centres. Second, a segmented land market can be imposed in both centres. Finally, we assume a mixed land market in the city, with a combination of a competitive land market in one centre and a segmented land market in the other. As an extension of the model used by Sivitanidou and Wheaton (1992) our model uses the same values for some of the exogenous parameters that they used in their simulation. The remaining variables that relate to the extensions of the model are selected in a way that facilitates comparison. The values are given in Table 3.3 below. The graphical and numerical presentations of the simulation output are given in Appendices 3AI and 3AII. The following section discusses the findings. In our model simulations we focus on two items: land market distortions, and investments in rail infrastructure.

Table 3.3: Value of exogenous parameters

Parameter	**Value**	**Parameter**	**Value**
Distance between nodes (b/2 mile)	10.00	National utility level u	1.20
Width of the city (mile)	1.00	Price of non-land consumption	1.00
β in utility function	0.50	Price of production output (p)	1.00
Agricultural rent ($/acre)	7500.00	Productivity per worker (prt)	22,371
Annual cost of transport		Other production costs (CE_1)	0
- Slow mode k_s ($)	350.00	Other production costs (CE_2)	0
- Fast mode k_f ($)	200.00	Cost advantage (CA) for centre 1 ($)	0
δ in floor space function	0.77	*Commercial land restrictions*	
Floor space per worker a_s (sq. ft.)	250.00	- Centre 1 (sq. miles)	1.80
Annual rent of capital ($/sq. ft.) r_c	7.00	- Centre 2 (sq. miles)	1.80

3.6.1 Effect of land market distortions

The competitive land market makes it possible for the production centres to acquire the required amount of land input for their production process at the competitive land rent. This leaves the relative cost of land and capital, and the capital to land ratio in the production technology of the centres unchanged. In contrast, in the segmented market, the limits on commercial land imposed in the centre(s) affect the commercial land rent and residential land rent. The restriction has a direct effect on the relative cost of capital and land. This in turn affects the wage-paying ability of the centres. Therefore, segmentation of the market has an effect on the production technology of the centres. The higher the commercial land restriction imposed on the centres, the higher the relative price of land to capital. Centres will then be increasingly capital-intensive in their production process. This leads to higher commercial land rents in the centres. In real life, this fact is visible in the form of high-rise buildings in the central areas of cities. On the other hand, the commercial land restriction reduces the wage-paying ability of the centres. Thus, at each location, residential land rents will be lower than at the corresponding locations under the competitive land market situation.

As Table 3AII.1 shows, the occurrence of restrictions on commercial land in the centres leads to an *increase* in total commercial land receipts *(0.277 versus 0.273)*: the increase in rent per unit dominates the *decrease* in area. This seems to indicate that restrictions on commercial land use improve opportunities to use land rents as a source of finance for infrastructure. However, the *increase* in commercial land rents is more than off-set by a *decrease* in residential land rents *(3.758 versus 3.786)*.

3.6.2 Effect of transportation investment

The main focus of this chapter is to determine the effect of investment in a fast mode on the urban economy. Given the open city assumption, the transport investments lead to a growth of the urban economy in terms of more residents (workers), a higher residential density, and a higher production level. The decrease in transportation costs causes an increase in demand for residential land and the numéraire good. This initially leads to a higher utility level. However, the potential higher utility level causes an inflow of new households into the city until the utility level is again equal to the national utility level. For an analysis of the benefits of such investments, our partial equilibrium model implies that welfare levels per household and profit levels remain unaffected in the long run. In a dynamic model, an initial increase in profits and income disposable for other consumption may be expected, but these increases will be gradually dampened by the arrival of new residents and new producers. In the long run, the only actor to benefit is the absentee owner who receives higher rents. Therefore, we focus on the effects of transport investments on land prices.

To trace the effect of the transportation investment, we compare three cases: namely 1) the base case (i.e. the unimodal case); 2) the partial bimodal case, where only one centre is served by the fast mode from two stations; and 3) the complete bimodal case, where the two centres are served by the fast mode from three stations. Investment in the fast transport mode makes commuting to the centres possible from a wider range of locations. Hence, the city size is enlarged by claiming land from agricultural use outside the city. In addition, the fast mode attracts denser residential settlements around the stations, which contributes to the population increase in the city. In general, the effect of the investment in fast transport has a different effect for the two land-market structures. In the competitive market, the investment does not affect the centres' level of wage-paying ability. However, the average residential land rents increase as a result of an increase in the density of settlement around the stations (compare the unimodal and the complete bimodal case in Tables 3AII.1 and 3AII.2 in Appendix 3AII). On the other hand, because of the unchanged wage level, the commercial land rent is not affected. The capital to land ratio that represents the technology of the centres remains the same. However, the size of the centres increases due to the increase in the number of employees arriving at the centres. In the segmented market case, some effects occur on both the residential and commercial sides. The additional employment induced by the fast mode has the same direction of effect on the residential and the commercial land rents in the city (as was discussed above). The important feature here is that, because of the increase in the demand of commercial land, coupled with the limited supply of land, the commercial land rent increases. This makes the non-labour input into the production process costly. Hence, the wage-paying ability of the centres declines. As a result, the average residential land rent declines compared with the case of the competitive market (1.34 versus 1.36).

We conclude from Table 3AII.1 that, under a distorted land market, the total commercial rent increase in the city, as a result of the introduction of rail, is higher than in a competitive market (*0.314-0.277* versus *0.302-0.273*). However, total residential rents decrease substantially due to segmentation, and thus the net effect on total rent receipts is clearly less favourable under distorted land markets than under competitive land markets. Hence, if capturing rents as a means to finance infrastructure is feasible at all locations, the competitive market offers the best opportunities. But, if these opportunities are only possible at commercial locations, the conclusion may change.

(a) Competitive position of centres

The effect of partial investment in the faster transport mode on the relative competitive position of the centres can be seen by allowing only one centre to benefit from such service. We can see this effect from Tables 3AII.1 and 3AII.2 (in Appendix 3AII) under the partial bimodal case. In the simulation, we allow a partial fast mode transport investment to serve only Centre 1. In general, the fast mode leads to the growth of city size and an increase in total rent in both market situations. However, it affects the average and total rent of the individual centres differently. In the competitive market, the investment leads to the decline of average and total residential land rent around the disadvantaged centre. In this case, even though the average commercial land rent is not affected, the total commercial land rent of the centre declines. In contrast, in the segmented market case, the effect of the investment leads to a decline for both average and total commercial and residential land rents. This shows that the advantaged centre grows at the expense of the disadvantaged centre by claiming more of the households residing in the area between the centres.

(b) Land rent loss or gain as a result of rail investment

The railway investments lead to an increase in the total land rents. However, this does not imply a uniform increase of rent levels everywhere in the city. In fact, there are places which experience a decline of rent levels. This phenomenon occurs in the segmented land market situation because investment in the railway also affects the wage level in the centres by altering the relative price of land to capital. Residential land rent is directly affected by the wage-paying ability at the centres. Compared with the baseline unimodal case, a partial railway investment leads to a wage increase in the centre which is not connected by rail and a decline in the wage level in the centre which is connected by rail. Thus, even though we observe an increase in the residential land rent levels around the newly introduced railway stations and the centre which has experienced a wage increase, rent levels around the centre which has a rail-connect decline because of the decline in the wage level. As we further expand the railway system by connecting both centres by rail, we see a decrease in the wage level in the newly connected centre due to an increase in labour supply. On the other hand, the wage level in the centre which was already connected by rail increases. This is because the supply of labour declines as a result of the commuting to the other centre made possible by the new rail investment. Thus, while residential land rent around the newly connected centre declines, the rent levels around the centre which was already connected by rail increase. Reverse effects are observed on the commercial land rents. As labour supply in the centres increases due to the investment in rail, commercial land rent increases in the case of segmented markets with fixed land supply (see Table 3AII.1 in Appendix 3AII).

(c) Effects of mixed land market

We can also allow a mixed land market for the two centres in the city and see what effect this has. So we assume a competitive land market for Centre 1 and a segmented land market for Centre 2. The simulated result is given in Table 3AII.2 in Appendix 3AII. Generally, as expected, the outcome is in-between the outcomes of the two uniform land market situations. In the unimodal and complete bimodal transport cases, alternating the land market situation between centres results in perfect symmetry. However, the partial bimodal transport case has some special features. Higher land rent receipts are achieved when the centre served by the railway has a competitive land market.

3.7 CONCLUSION

Generally, investment in the fast (rail) transport mode results in city growth, in terms of both area size and population, an increase in rent receipts, and denser residential settlements. However, the effect of the investment for individual centres and their corresponding residential areas depends on the underlying land market conditions. As investments in railways steadily increase from a unimodal to a complete bimodal situation, rent-losing and gaining phenomena are observed along segments of the city in the segmented land market situation.

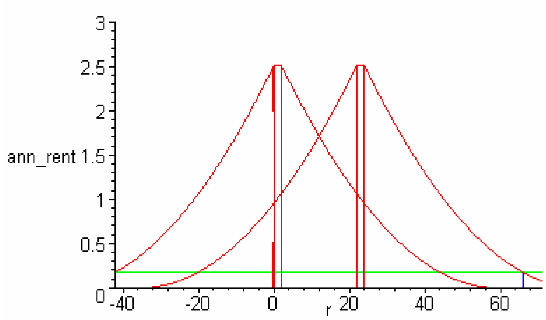
When land rents are captured as a source of investment for railway infrastructure, the increase in total rents is highest in the competitive land market situation. But it is important to realize that the rent increases are spread widely in the urban metropolitan area, which may make them difficult to collect in real-world situations. Of course, the most focussed rent increases take place near the railway stations. Of special importance is the finding that, in the case of segmented markets, the total commercial rent receipts are higher than in a competitive land market situation. Thus, as long as rent capturing is limited to commercial land use, the segmented land market is not as unfavourable as one might expect. The issue of land market distortions is important because these distortions may have decisive impacts on long-run changes in transport demand in response to changes in the transport system. In the partial bimodal plan, connecting the centre under a competitive land market results in higher land market receipts compared with connecting the centre under a segmented land market.

In this chapter we have analysed the impact of a second transport mode on the dynamics of centres in a metropolitan area, under the assumption that the additional infrastructure may reinforce or weaken the existing commercial centres. However, transport investments may also have far-reaching effects on spatial structure, since they may stimulate the emergence of new centres. This theme of new centre formation has not been addressed in the present chapter. Instead, we have focussed explicitly on the demand for commercial land and the implications of distortions for the land market. Analysing the possible emergence of additional centres falls outside the scope of this model, but is certainly a promising extension. In order to achieve this aim, the model should be developed in the direction of a more explicit treatment of production processes and agglomeration economies.

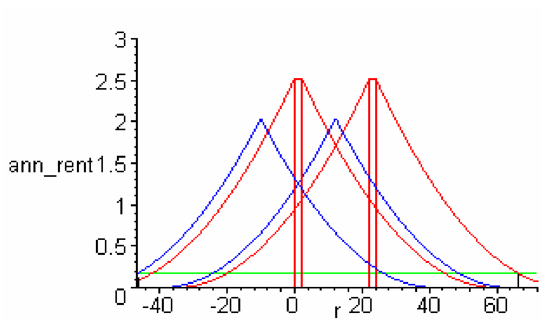
The discussion that transport nodes in an urban area are faced with a downward-sloping rent gradient is used as a basis for the empirical discussion addressed in the following chapter. In the following chapter we discuss the effect of railway accessibility on residential house prices. Railway accessibility is explained by both the distance to the railway stations and the service levels provided at the station.

APPENDIX 3AI: Graphical presentation of simulation results

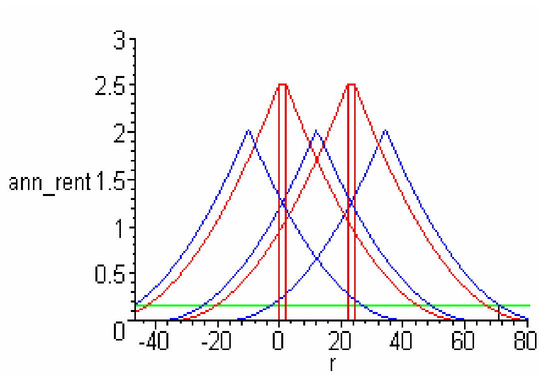
Figure 3AI. 1: Competitive market



a) Rent curves for the unimodal-bicentric city case: a competitive market situation

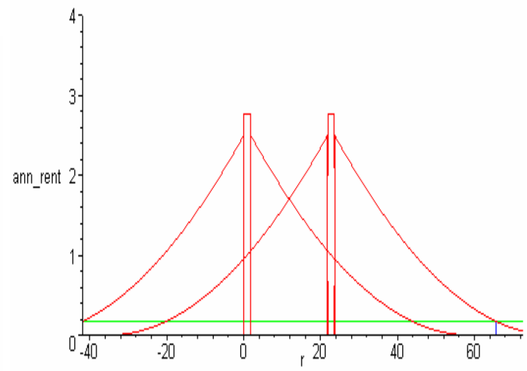


b) Rent curves for the partial bimodal city case: a competitive market situation

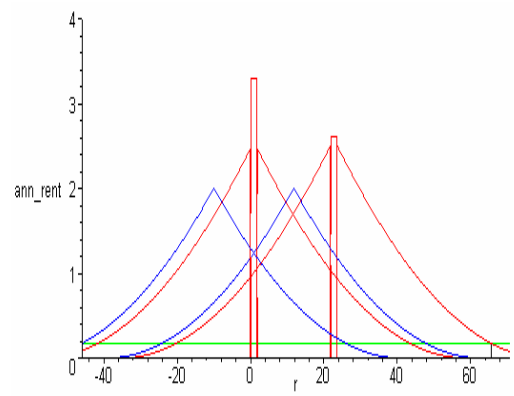


c) Rent curves for the complete bimodal city case: a competitive market situation

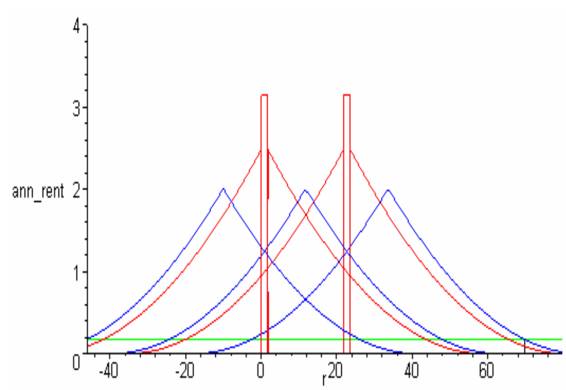
Figure 3AI. 2: Segmented market



a) Rent curves for the unimodal-bicentric city case: a segmented market situation



b) Rent curves for the partial bimodal city case: a segmented market situation



c) Rent curves for the fully improved bimodal city case: a segmented market situation

APPENDIX 3AII: Numerical presentation of simulation output

Table 3AII. 1: Simulation output for both markets

CASES		Centre	Fast mode	Exogenous Para. Comm. Land Area sq. miles	Annual Wages ($)	Endogenous Variable Resid. Land sq. miles	Labour Supply	Comm. Land sq. miles	Capital Land Ratio	Annual Resid. Rent at Edge of Centre $/sq. ft.	Annual Com. Rent at Edge of Centre $/sq. ft.	Aver. Annual Resid. Rent $/sq. ft.	Aver. Annual Com. Rent $/sq. ft.	Annual Floor Space Rent $/sq. ft.	Total Resid. Rent Bill. $	Total Com. Rent Bill. $	Total Rent (Bill. $)
Unimodal	**Competitive market**	1	no		20,000	52.17	250,000	1.95	1.20	2.51	2.51	1.30	2.51	9.48	1.893	0.136	2.029
		2	no		20,000	52.17	250,000	1.95	1.20	2.51	2.51	1.30	2.51	9.48	1.893	0.136	2.029
		Total/ average				**104.34**	**500,000**	**3.90**				**1.30**	**2.51**		**3.786**	**0.273**	**4.058**
	Segmented market	1	no	1.80	19,947	52.02	248,751	1.80	1.32	2.50	2.77	1.30	2.77	9.70	1.879	0.139	2.018
		2	no	1.80	19,947	52.02	248,751	1.80	1.32	2.50	2.77	1.30	2.77	9.70	1.879	0.139	2.018
		Total/ average				**104.04**	**497,502**	**3.60**				**1.30**	**2.77**		**3.758**	**0.277**	**4.035**
Partial bimodal	**Competitive market**	1	yes		20,000	58.60	290,195	2.26	1.20	2.51	2.51	1.38	2.51	9.48	2.252	0.158	2.410
		2	no		20,000	50.03	237,352	1.85	1.20	2.51	2.51	1.28	2.51	9.48	1.786	0.129	1.916
		Total/ average				**108.63**	**527,547**	**4.11**				**1.34**	**2.51**		**4.038**	**0.288**	**4.326**
	Segmented market	1	yes	1.80	19,846	57.97	285,107	1.80	1.58	2.47	3.30	1.36	3.30	10.10	2.196	0.166	2.361
		2	no	1.80	19,979	50.16	238,003	1.80	1.25	2.51	2.61	1.28	2.61	9.57	1.791	0.131	1.922
		Total/ average				**108.13**	**523,110**	**3.60**				**1.34**	**2.99**		**3.986**	**0.297**	**4.283**
Complete bimodal	**Competitive market**	1	yes		20,000	56.46	275,860	2.15	1.20	2.51	2.51	1.36	2.51	9.48	2.135	0.151	2.287
		2	yes		20,000	56.46	275,860	2.15	1.20	2.51	2.51	1.36	2.51	9.48	2.135	0.151	2.287
		Total/ average				**112.92**	**551,720**	**4.30**				**1.36**	**2.51**		**4.271**	**0.302**	**4.573**
	Segmented market	1	yes	1.80	19,876	56.10	273,894	1.80	1.50	2.48	3.14	1.34	3.14	9.98	2.100	0.157	2.257
		2	yes	1.80	19,876	56.10	273,894	1.80	1.50	2.48	3.14	1.34	3.14	9.98	2.100	0.157	2.257
		Total/ average				**112.21**	**547,788**	**3.60**				**1.34**	**3.14**		**4.199**	**0.314**	**4.514**

Table 3AII. 2: Simulation output for both the three transport cases with mixed land markets between the centres

CASES		Centre	Fast mode	Exogenous Para. Comm. Land Area sq. miles	Exogenous Para. Annual Wages ($)	Endogenous Variable Resid. Land sq. miles	Labour Supply	Comm. Land sq. miles	Capital Land Ratio	Annual Resid. Rent at Edge of Centre $/sq. ft.	Annual Com. Rent at Edge of Centre $/sq. ft.	Aver. Annual Resid. Rent $/sq. ft.	Aver. Annual Com. Rent $/sq. ft.	Annual Floor Space Rent $/sq. ft.	Total Resid. Rent Bill. $	Total Com. Rent Bill. $	Total Rent (Bill. $)
Unimodal	**Competitive market**	1	no		20,000	52.17	250,002	1.95	1.20	2.51	2.51	1.30	2.51	9.48	1.893	0.136	2.029
	Segmented market	2	no	1.80	19,947	52.02	248,751	1.80	1.32	2.50	2.77	1.31	2.77	9.70	1.879	0.139	2.018
		Total/ average				**104.19**	**498,753**	**3.75**				**1.31**	**2.64**		**3.772**	**0.275**	**4.047**
	Segmented market	1	no	1.80	19,947	52.02	248,752	1.80	1.32	2.50	2.77	1.30	2.77	9.70	1.879	0.139	2.018
	Competitive market	2	no		20,000	52.17	250,001	1.95	1.20	2.51	2.51	1.29	2.51	9.48	1.893	0.136	2.029
		Total/ average				**104.19**	**498,753**	**3.75**				**1.30**	**2.64**		**3.772**	**0.275**	**4.047**
Partial bimodal	**Competitive market**	1	yes		20,000	58.63	290,350	2.26	1.20	2.51	2.51	1.38	2.51	9.48	2.253	0.158	2.412
	Segmented market	2	no	1.80	19,982	49.95	236,792	1.80	1.24	2.51	2.60	1.28	2.60	9.56	1.780	0.130	1.910
		Total/ average				**108.58**	**527,142**	**4.06**				**1.33**	**2.55**		**4.034**	**0.288**	**4.322**
	Segmented market	1	yes	1.80	19,846	57.94	284,938	1.80	1.58	2.47	3.30	1.36	3.30	10.10	2.194	0.165	2.360
	Competitive market	2	no		20,000	50.25	238,674	1.86	1.20	2.51	2.51	1.29	2.51	9.48	1.798	0.130	1.928
		Total/ average				**108.19**	**523,612**	**3.66**				**1.33**	**2.90**		**3.992**	**0.296**	**4.287**
Complete bimodal	**Competitive market**	1	yes		20,000	56.46	276,985	2.16	1.20	2.51	2.51	1.34	2.51	9.48	2.249	0.151	2.400
	Segmented market	2	yes	1.80	19,876	56.10	273,985	1.80	1.50	2.48	3.14	1.35	3.14	9.98	1.985	0.157	2.142
		Total/ average				**112.56**	**550,970**	**3.96**				**1.34**	**2.80**		**4.234**	**0.308**	**4.542**
	Segmented market	1	yes	1.80	19,876	56.10	273,985	1.80	1.50	2.48	3.14	1.35	3.14	9.98	1.985	0.157	2.142
	Competitive market	2	yes		20,000	56.46	276,985	2.16	1.20	2.51	2.51	1.34	2.51	9.48	2.249	0.151	2.400
		Total/ average				**112.56**	**550,970**	**3.96**				**1.34**	**2.80**		**4.234**	**0.308**	**4.542**

Chapter 4

4 The impact of rail transport on house prices: an empirical analysis of the Dutch housing market[8]

4.1 INTRODUCTION

Hedonic pricing methods explain the value of real estate in terms of the features of the property. This approach treats a certain property as a composite of characteristics to which value can be attached. The sum of the value of the individual characteristics makes up the value of the property as a whole. Studies on real estate prices generally categorize the value-bearing features of properties into three types: namely, physical, accessibility and environmental (Fujita 1989; Bowes and Ihlanfeldt 2001). Several studies have been conducted that focus on different features of interest. Accessibility, as provided by different modes of transportation and railways in particular has also received attention. In order to single out the effect of railway stations on property values, it is suggested in the literature that stations should be seen as nodes in a transport network and places in an area (Bertolini and Spit 1998). Based on this framework, recent empirical studies treat the node feature and the place feature of a station separately. The former characteristic accounts for the accessibility effect, which is generally positive. The latter feature accounts for externalities of the station and can have both positive and negative effects. Bowes and Ihlanfeldt (2001) pointed at the retail employment and crime that stations attract in addition to the accessibility feature of a station. By including the three categories of property features mentioned above, this chapter examines the effect of railway stations on Dutch house prices. There are three types of rail service in the Netherlands: light rail services (trams); heavy rail services (metro lines); and commuter rail services. The services of the first two are limited within the main cities. However, the third type serves the whole country. This chapter assesses the effect of accessibility provided by these commuter railway stations on house prices in the area. The main focus of this chapter is the analysis of the impact of railway accessibility on residential house prices. However, as Voith (1993) pointed out, highway accessibility is an important competitor to rail

[8] This chapter is based on Debrezion, Ghebreegziabiher, Eric Pels and Piet Rietveld (2006). "Impact of railway station on Dutch residential housing market". Tinbergen Institute discussion paper TI 2006-031/3.

accessibility: “The presence of other facilities that increase accessibility like highways, sewer services and other facilities influence the impact area in the same fashion.” The benefits of these facilities and services are also capitalized in urban property values (Damm et al. 1980). Thus, to single out the effect of railway accessibility, highway accessibility is represented in our analysis by means of distance to points of highway entry and exits.

The accessibility and nuisance effects of a railway station are functions of distance between the station and the house under consideration. As the distance increases, the impact of both these effects on the house price declines. The level of accessibility at a railway station is measured by the quality of the railway network, which can be defined in terms of: the number of destinations that can be reached from the station, the frequency of services at the station, and other departure-station-related facilities. Stations with higher network quality (i.e. a larger number of destinations and a higher frequency of trains) have a higher accessibility index, and are expected to have a relatively high positive effect on the house prices. Railway stations at the same time impose localized negative environmental effects on house prices due to noise nuisance. An important difference between the two effects is that the accessibility effects are concentrated around nodes (railway stations), whereas the negative noise effects take place everywhere along the railway line.

In this chapter we determine the impact of the three railway features: namely, railway station proximity; rail service levels; and proximity to the railway line, on the prices of residential properties. The data for the analysis in this chapter includes the sales and prices of residential properties in the Netherlands. As a result of the transportation cost and time savings made possible, households are expected to be willing to pay higher prices for living close to the station compared with other locations. This is because the commuting (time) costs are relatively low for people living near a station. Furthermore, leisure activities that involve rail transport are more accessible. This chapter only covers the sales of residential properties. In Chapter 8 we study the effect of the railway station on commercial property values.

4.2 LITERATURE REVIEW

In general, the empirical studies conducted in this area are diverse in methodology and focus. Although the functional forms can differ from study to study, the most common methodology encountered in the literature is hedonic pricing. However, no consistent relationship between proximity to railway stations and property values is recorded. Furthermore, the magnitudes of these effects can be minor or major. One of the earliest studies, Dewees (1976), analysed the relationship between travel costs by railway and residential property values. Dewees found that a subway station increases the site rent perpendicular to the facility within $\frac{1}{3}$ mile of the station. Similar findings confirmed that the distance of a lot from the nearest station has a statistically significant effect on the property value of the land (Damm et al. 1980). Consistent with these conclusions, Grass (1992) later found a direct relationship between the distance of the newly opened metro and residential property values. Some of the extensively studied metro stations in the U.S., though ranging from small to modest impact, show that properties close to the station have a higher value than properties farther away (Giuliano 1986; Bajic 1983; Voith 1991). However, there are studies which have also found insignificant effects (Lee 1973; Gatzlaff and Smith 1993). Evidence from other studies indicates little impact in the absence of favourable factors (Gordon and Richardson 1989; Giuliano 1986). For a detailed documentation of the findings, we refer to Vessali 1996; Smith and Huang 1995; NEORail II 2001; and GVA Grimley 2004. In general, some studies indicate a decline in the historical impact of railway stations on property values. This is attributed to improvements in accessibility, advances in telecommunications, computer networks, and other areas of technology that were said to make companies "footloose" in their location choices (Gatzlaff and Smith 1993).

The impact of railway stations on property values varies as a result of several factors. First, railway stations differ from each other in terms of the level of service provided, explained in terms of frequency of service, network connectivity, service coverage, etc. The meta-analysis in Debrezion et al. (2006) (see Chapter 2) shows that different types of railway stations have different levels of impact on property value. Commuter railways have a relatively high impact on property values (Debrezion et al. 2006; Cervero and Duncan 2001; NEORail II 2001; Cervero 1984). Railway stations also differ in the level and quality of facilities. Stations with a higher level and quality of facilities are expected to have greater impact on the surrounding properties. The presence and number of parking lots is one of the many station facilities that have received attention in the literature. Bowes and Ihlanfeldt (2001) found that stations with parking facilities have a higher positive impact on property values. In addition, the impact a railway station produces depends on its proximity to the CBD. Stations which lie close to the CBD produce a greater positive impact on the property value (Bowes and Ihlanfeldt 2001). In another study, Gatzlaff and Smith (1993) have claimed that the variation in the findings of the empirical work is attributed to local factors in each city.

Second, railway stations affect residential and commercial properties differently. Most studies have treated the effect of railway stations on the different property types separately. The range of the impact area of railway stations is larger for residential properties, whereas the impact of a railway station on commercial properties is limited to immediately adjacent areas. Generally, it has been shown that the impact of railway stations on commercial properties is greater than the impact on residential properties within a short distance of the stations (Cervero and Duncan 2001; Weinstein and Clower 1999). This finding is in line with the assertion that, railway stations as focal and gathering points attract commercial activities, which increase commercial property values. However, contrary to this assertion, Landis et al. (1995) determined a negative effect on commercial property values.

Third, the impact of railway stations on property values is subject to the demographic segmentation of neighbourhoods. Income and social (racial) divisions are common. Proximity to a railway station is of higher value to low-income residential neighbourhoods than to high-income residential neighbourhoods (Nelson 1998; Bowes and Ihlanfeldt 2001). The reason is that low-income residents tend to rely more on public transport, and thus attach a higher value to living close to the station. Because of the fact that this group of people mostly depend on slow modes (walking and bicycle) to access the stations, locations adjacent to railway station are expected to constitute poor segments. On the other hand, the high population movement in the immediate location gives rise to the development of retail activities which eventually increase the value of commercial properties, but it may at the same time attract criminality (Bowes and Ihlanfeldt 2001). Bowes and Ihlanfeldt outlined that a significant relation was observed between stations and crime rates. However, no proximity variable shows a significant effect on retail employment. In this model, the immediate neighbourhood is affected by the negative impact of the station. Thus, the most immediate properties (within ¼ mile of the station) were found to have an 18.7% lower value. Properties that are situated between one and three miles from the station, are however, are more valuable than those further away.

4.3 DATA AND DESCRIPTIVES

(A) HOUSE CHARACTERISTICS

The data used in the analysis of this chapter covers sales transactions of the Dutch residential housing market for a period of 17 years from 1985 to 2001. These transactions are recorded by the Dutch Brokers Association (NVM). The data incorporate information related to the price of the dwellings, the characteristics of the dwellings and some environmental features. To further enrich the data set, each of the houses sold is geo-coded separately to enable us to compute the distances to the railway stations and highway entry/exit points. Some houses are geo-coded at the precise house address level, and the rest are geo-coded at the 6-digit (e.g. 1234XX) postcode level, which is an area comprising up to about 20 houses. Apart from the house characteristics, a number of accessibility and neighbourhood features are used. The land use data were acquired from the Central Office of Statistics for the Netherlands (the CBS). These data are available at the 4-digit postcode level. Moreover, population-related data are available at this level of aggregation. Income levels of the population in the postcode area, the density and population composition, in particular the share of foreigners in the area, are used in our analysis.

The accessibility data relate to two transport modes: railway and highway. The locations of all railway stations and highway entry/exit points are identified. The distance from the houses to these points was determined by GIS methods. The distance to the nearest highway entry/exit points is expected to account for the car-based accessibility. This chapter uses two references for a railway station: the nearest railway station, and the most frequently-chosen railway station. The nearest station is easily determined using GIS methods. The identification of the most frequently-chosen station was based on the survey study of the Dutch National Railway Company (NS). It is given at the 4-digit postcode area level.

In Table 4.1 some descriptive statistics on the three categories of factors affecting property values are given. For the physical features of the houses we use a large number of relevant items. Examples are the surface area of the house (that includes the built-up and non-built up part of the property), age of the house, the number of rooms and number of bathrooms; all these variables are continuous. The rest of the physical characteristics, such as the monument status of the dwelling, the availability of a gas heater, the presence of an open fireplace, the presence of a garden and a garage are indicated by dummy variables. The mean values for some of these features are given in Table 4.1. The descriptive statistics are based on 663,024 houses sold in the time period considered. The features in the accessibility category include distance to the railway station, the frequency of trains, and the distance to the nearest highway entry/exit point (both with respect to the most frequently-chosen station for residents in the postcode area and the nearest station to the house). The analysis also includes the perpendicular distance to railway lines in an effort to capture the noise effect of railways.

Table 4.1: Descriptive statistics of house characteristics

	Minimum	Maximum	Mean	Std. Deviation
Dependent variable				
Transaction price in euros	9076	5,558,800	123,187	95,678
Independent variables				
1. House features				
Surface area in sq. metres	11	99,998	443	1890
Building age in years	0	996	38	40
Total number of rooms	1	39	4.47	1.34
Number of bathrooms	0	4	0.87	0.58
Dummy variables				
Monument status			0.009	
Gas heater			0.136	
Open fireplace			0.186	
Garage			0.335	
Garden			0.783	
2. Accessibility features				
Distance to nearest railway station (m)	3	25,498	3,486	3441
Distance to most frequently-chosen railway station (m)	10	35,643	4,245	5064
Frequency (trains/day at the most frequently-chosen station)	18	788	268	217
Frequency (at the nearest station)	18	788	169	151
Distance to highway entry/exit (m)	0	39,541	3,978	4711
3. *Nuisance feature*				
Distance to railway line	0	23,696	2,351	3,052
4. Environmental features				
Household income in euros (4-digit postcode level)	3136	26200	11480	1805
Population composition (percentage of foreigners)	0.010	.890	.642	0.918

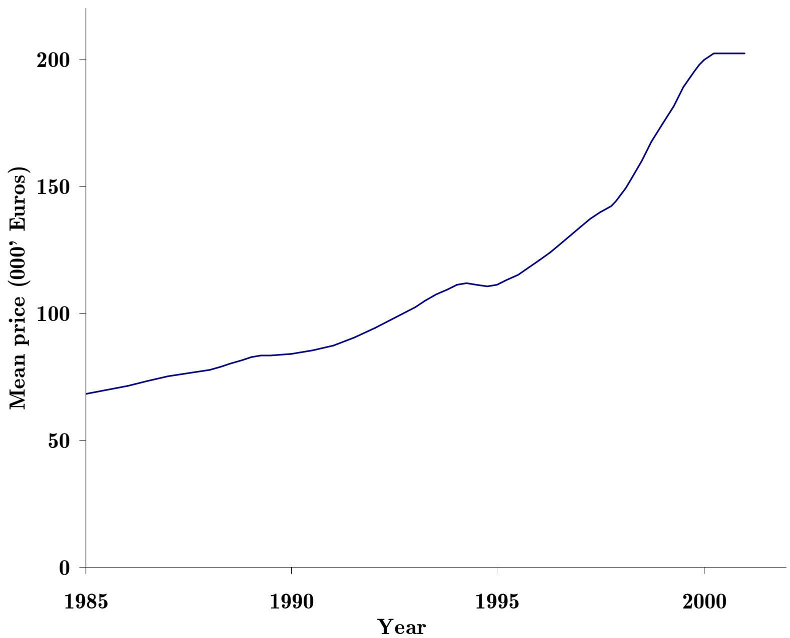


Figure 4.1: Mean price of houses by year

The distance to the most frequently-chosen station is on average about 1 kilometre longer than the average distance to the nearest railway station. The average frequency of trains at the most frequently-chosen station is more than 100 trains per day over the average frequency of trains at the nearest railway station. This gives an indication of the trade-off travellers make between proximity of stations and the level of service they offer. Figure 4.1 shows the average transaction price in each year. This increase can be attributed to the combined effect of inflation and real value increase.

(B) RAILWAY STATION CHARACTERISTICS

The data of particular interest in this study concerns railway accessibility and associated noise or congestion. Railway accessibility can be explained by two features: the proximity feature, and service level features. The first feature is more or less captured by the distance measure, whereas various features can contribute to the service level. Examples include the number of trains leaving the station per time unit, and network connectivity as measured by the number of destinations served by the station. In addition, service level may also include facilities that supplement railway transport. For example, the availability of parking space, the park-and-ride status of the station and the availability of bicycle storage can be mentioned. The overall

Dutch railway network is composed of about 360 stations. Our data allows us to use the most frequently-chosen departure station for households aggregated at the 4-digit postcode level.

Table 4.2: Descriptive statistics for the railway station characteristics

	No. stations	Minimum	Maximum	Mean	Std. Deviation
Rail service					
Frequency of trains per day		18	788	113	103
Destinations reached without a transfer		1	114	16	14
Destination reached with one transfer		8	246	87	53
Travel demand					
Total passenger turnout per day		46	145,700	5,600	13,770
Station type					
Intercity stations	64			0.18	
Station Facilities (dummy variables)					
Train taxi	109			0.30	
Bicycle stand	96			0.27	
Bicycle storage	264			0.74	
Bicycle rent	114			0.31	
Park-and-ride	49			0.14	
Parking	326			0.91	
Taxi	163			0.45	
Car rent	1			0.00	
Luggage deposit	64			0.18	
International connection	22			0.06	

4.4 METHODOLOGY

The hedonic pricing methodology is found to be effective in singling out the effect of one characteristic from a number of characteristics of which a property is composed (Rosen 1974). This chapter uses this approach to determine the effect of the three categories of house features in general, and railway accessibility in particular. A semi-logarithmic specification is adapted. Thus, the dependent variable in our analysis is the natural logarithm of the transaction price of residential houses. A wide range of independent variables that are expected to explain the house prices are included. These include the physical characteristics of the houses, environmental amenities, and the accessibility variables that correspond to the houses under study. Because the data set covers a relatively long period, and house prices have increased continuously during the last decade, temporal effects are also expected to play a role in explaining the variation in the sales price of houses. Thus, we include sales year dummies to capture the temporal effects. These account for the inflation, real value changes, and other temporal effects across the time period. To account for the spatial effect, regional dummies are included at the municipality level. The main focus of the analysis here is the effect of railway station proximity and service quality of the stations. We also include the effect of proximity to highway entry/exit points in order to account for competition by the car.

MODEL SPECIFICATION

Even though the data include a longer period, we could not organize our data in a panel structure because there were not many repeated sales over the time. Therefore, our data is organized in a cross-sectional pattern. The semi-logarithmic hedonic specification is widely used in the property value literature. Its use is motivated by the fact that it gives robust estimates and it enables convenient coefficient interpretation. The general structure of the model we adopt here is:

$$Ln(P_i) = B'_0 + B'_1 X_{i1} + B'_2 X_{i2} + \ldots + B'_n X_{in} + \varepsilon_i, \tag{1}$$

where, P_i is the price house i, and $X_{i1} \ldots X_{in}$ are vectors of explanatory variables for the price of house i. The dependent variable is given in the natural logarithmic form; thus, the values of the coefficients represent percentage change. The specifications used in the estimations are given by Equations 2 and 3. Distances from the houses to the railway station and line and highway entry /exit points are classified according to several distance categories. The first model includes the distance and frequency effect (station quality) separately. The second model includes the interaction between distance and frequency. In both specifications, proximity to the railway station and the railway line are treated in piecewise fashion. Frequency of trains at the reference station is given in continuous form. The models have the following form:

$$\begin{aligned} \ln(\text{tranPrice}_i) = \alpha &+ \beta'_{HC} \times HouseChr_i + \beta'_{dc} \times Distcategrail_i + \beta_{freq} \times \ln(FreqT_i) \\ &+ \beta'_{hw} \times Distcateghway_i + \beta'_{railline} \times Drailline_i + \beta'_{Neighb} \times Neighb_i \\ &+ \beta'_{\mathrm{Re}gion} \times Dregional_i + \beta'_{time} \times Dtime_i + \varepsilon_i; \end{aligned} \tag{2}$$

$$\begin{aligned} \ln(\text{tranPrice}_i) = \alpha &+ \beta'_{HC} \times HouseChr_i + \beta'_{DC \otimes Freq} \times Distcategrail_i \otimes \ln(FreqT_i) \\ &+ \beta'_{hw} \times Distcateghway_i + \beta'_{railline} \times Drailline_i + \beta'_{Neighb} \times Neighb_i\,. \\ &+ \beta'_{\mathrm{Re}gion} \times Dregional_i + \beta'_{time} \times Dtime_i + \varepsilon_i\,. \end{aligned} \tag{3}$$

where, tranPrice_i represents the transaction price of house i; $HouseChr_i$ is a vector of house characteristics for house i, which includes variables for type of house, surface area, total number of rooms, number of bathrooms, presence of garage and garden for the house, presence of gas heater and open fireplace, monument status, age of the building; $Distcategrail_i$ is a vector of dummy variables representing the distance category at which house i is located from a station. To see the smoothness of the effect, we use categories with a 500 metres range except in the two inner circle categories of the station, which are 250 metres each. Thus, we have 31 categories of distances up to 15 kilometres. Areas beyond this limit are taken as a reference group in the estimation; $FreqT_i$ is the frequency of trains at the station to which the distance is computed and is given in trains per day. In our analysis we make two station considerations: the nearest vs the most frequently-chosen station in the post-code area; $\otimes$ is the Kronecker product to indicate the cross-product of distance classes and frequency of trains at the reference station; $Distcateghway_i$ is a vector of dummies representing the distance category at which a highway entry/exit point is located from the house. In the same fashion as the railway distance categories, we also have 31 distance categories for these variables; $Drailline_i$ is a vector of two dummy variables representing the distance category in which the house is situated in relation to the railway line. This is expected to account for the noise effect of trains. The railway noise is expected to have a localized effect and thus we compare the effect of noise on two nearby distance categories against the rest; $Neighb_i$ is a vector of neighbourhood characteristics including income, ratio of foreigners and share of land use types. It is given at the 4-digit postcode level; $Dregional_i$ is a vector of dummy variables representing the municipality to which the house belongs; $Dtime_i$ is a vector of time dummy variables representing the year when the transaction took place; and ε_i is the error term.

All in all, the total number of explanatory variables in the hedonic pricing models is 344. Of these, 34 relate to house characteristics, 28 to neighbourhood features, 16 to time series dummies, and 203 to municipality dummies. The remaining 63 variables represent railway and highway accessibility. In the presentation of the estimations below, we focus on the impact of the accessibility variables. The municipality dummies can be considered to represent the many municipality-specific factors that may affect house values. Thus, the effects we find for railway station proximity have been corrected for municipality-specific impacts.

Generally, the price of houses is expected to rise as the distance to the railway station and/or highway entry/exit points decreases. At the same time, the influence of a station on the house prices is expected to increase with an increase in the service level provided by the station, as given by frequency of trains and the number of destinations directly served by the station. However, the latter two variables are highly correlated, so we prefer to include one of the two in our estimation. We find the frequency variable more telling since it addresses scheduling and waiting time aspects, an important dimension of generalized costs. In addition, frequency is related to reliability since delays are less disturbing in the case of high frequency.

4.5 ESTIMATION RESULTS AND DISCUSSIONS

Table 4.3 gives four estimation results based on Equations 2 and 3. To save space, we only report the coefficients of the factors that relate to railway aspects. The complete estimation results are available upon request from the author. The first two estimations correspond to the simple linear effect of piecewise distances and the frequency-of-trains effect treated separately, as given by Equation 2. The last two estimations are based on the model given by Equation 3. The cross-distance frequency estimation gives the effect of frequency of trains on house prices for each of the distance classes. The semi-log nature of the model makes the interpretation of the coefficients easier. Each coefficient for the distance categories in the first two estimations shows that the percentage effect on house prices of those distances to the station compared with houses located beyond 15 kilometres. Thus, we observe a difference as big as 32% in house prices for houses within 500m of the nearest station and houses located more than 15 kilometres from the stations. This difference gets smaller in the case of the most frequently-chosen station effect (about 27%), where we encounter the peak house price to be between 250 and 500 metres. The trend of the effect sizes for this specification is given in Figure 4.2. This figure shows irregularity in the distance category of 7.5 to 8 kilometres. This is due to the small number of observations in this category. Such irregularities are inevitable when small distance classes are used. The difference between the distance effect of the nearest and most frequently-chosen station is remarkable. The advantage of being close to the station is not so large in the case of the most frequently-chosen station compared with the nearest station. The reason is that the most frequently-chosen station apparently has extra qualities that make it more attractive than the nearest station. Hence, it may be expected that distance to the station matters less in the price effect on real estate. The mirror image is that the quality of the station, as reflected by, amongst other things, the frequency, has a larger effect. This explains why the frequency elasticity in Table 4.3 is so high for the most frequently-chosen station compared with the nearest station (0.09 versus 0.03). A doubling of frequency of trains at the most frequently-chosen station results in a 9% house price increase in the postcode area compared with a 3% increase for the case of the nearest railway station (see the first two columns of Table 4.3). Finally, we find clear negative effects of railway noise on house values: houses located in the zone within 250 metres from a railway line are about 5% less expensive than houses located at 500m or more. For the zone between 250-500m, intermediate values are found.

However, the measure of the frequency-of-trains effect discussed above is crude since it is not distance dependent. The point is that a frequency increase is probably more important for dwellings close to a station than it is for dwellings far away. The last two columns of Table 4.3 provide the estimation of the cross-distance-frequency effect. Doubling the frequency of trains in the nearest station results in as much as a 3.5% price increase for houses located up to 2 kilometres away compared with the effect on dwellings located beyond 15 kilometres. On the other hand, doubling the frequency of trains at the most frequently-chosen station results in a price increase of about 3.0% for the same distance category. The pattern in the elasticities of frequency for the different distance categories is depicted graphically in Figure 4.3. These estimations demonstrate that the value of property may depend on the proximity to more than one railway station. We will not investigate this issue in more detail here, but this is an indication that railway station accessibility is a more complex concept than one might think: it involves competition between railway stations.

Furthermore, the percentage effect of different levels of frequency is given in Table 4.4 below. The table shows – not surprisingly – that the effect of railway proximity is largest in the case of a station with a high level of service. Note that such a differentiated effect is not present in the specification given by Equation 2. However, the frequency impact is smaller than one might expect. Nevertheless, the price curves are clearly steeper around stations with higher frequencies. Further, we find that, even for stations with a small number of trains, a substantial effect of railway presence is found. Note that this estimation is based on a specification where corrections were carried out for a large number of other variables. In particular, a dummy has been added for each municipality so that it has been assured that the results found do not capture the effects of other variables such as population density or other municipality-specific factors.

Table 4.3: Estimation of railway station effects on house values: piecewise distance effect
(N.B. Only railway-related parameters are presented.)

Variable	Nearest Station		Most frequently chosen station		Cross distance-frequency of trains effect			
					Nearest Station		Most frequently chosen Station	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
(Constant)	8.966***	0.009	8.775***	0.009	9.189***	0.008	9.232***	0.008
raildist250	0.323***	0.006	0.271***	0.004	0.050***	0.001	0.043***	0.001
raildist250_500	0.321***	0.005	0.274***	0.003	0.050***	0.001	0.044***	0.001
raildist500_1000	0.315***	0.005	0.260***	0.003	0.049***	0.001	0.043***	0.001
raildist1000_1500	0.308***	0.005	0.246***	0.003	0.048***	0.001	0.042***	0.001
raildist1500_2000	0.316***	0.005	0.245***	0.003	0.049***	0.001	0.043***	0.001
raildist2000_2500	0.296***	0.005	0.232***	0.003	0.045***	0.001	0.041***	0.001
raildist2500_3000	0.287***	0.005	0.203***	0.003	0.042***	0.001	0.036***	0.001
raildist3000_3500	0.277***	0.005	0.203***	0.003	0.041***	0.001	0.038***	0.001
raildist3500_4000	0.299***	0.005	0.201***	0.003	0.046***	0.001	0.038***	0.001
raildist4000_4500	0.284***	0.005	0.181***	0.003	0.042***	0.001	0.035***	0.001
raildist4500_5000	0.252***	0.005	0.160***	0.003	0.037***	0.001	0.033***	0.001
raildist5000_5500	0.238***	0.005	0.153***	0.003	0.033***	0.001	0.033***	0.001
raildist5500_6000	0.234***	0.005	0.133***	0.004	0.033***	0.001	0.030***	0.001
raildist6000_6500	0.226***	0.006	0.106***	0.004	0.031***	0.001	0.027***	0.001
raildist6500_7000	0.229***	0.006	0.105***	0.004	0.032***	0.001	0.028***	0.001
raildist7000_7500	0.204***	0.006	0.093***	0.004	0.027***	0.001	0.026***	0.001
raildist7500_8000	0.235***	0.006	0.006***	0.004	0.034***	0.001	0.009***	0.001
raildist8000_8500	0.215***	0.006	0.065***	0.004	0.029***	0.001	0.021***	0.001
raildist8500_9000	0.266***	0.006	0.098***	0.004	0.040***	0.001	0.028***	0.001
raildist9000_9500	0.213***	0.007	0.106***	0.004	0.029***	0.001	0.030***	0.001
raildist9500_10000	0.177***	0.007	0.100***	0.004	0.023***	0.001	0.028***	0.001
raildist10000_10500	0.158***	0.007	0.047***	0.005	0.019***	0.001	0.018***	0.001
raildist10500_11000	0.069***	0.007	0.040***	0.005	0.002	0.001	0.017***	0.001
raildist11000_11500	0.037***	0.008	0.038***	0.005	-0.005***	0.002	0.016***	0.001
raildist11500_12000	0.036***	0.008	0.053***	0.005	-0.006***	0.002	0.022***	0.001
raildist12000_12500	0.036***	0.009	0.070***	0.005	-0.005***	0.002	0.026***	0.001
raildist12500_13000	0.022***	0.009	0.070***	0.005	-0.011***	0.002	0.024***	0.001
raildist13000_13500	0.007	0.009	0.047***	0.005	-0.013***	0.002	0.020***	0.001
raildist13500_14000	0.028***	0.008	0.034***	0.005	-0.007***	0.002	0.016***	0.001
raildist14000_14500	0.031***	0.008	0.062***	0.005	-0.003	0.002	0.021***	0.001
raildist14500_15000	0.029***	0.009	0.035***	0.005	-0.002	0.002	0.015***	0.001
Log (frequency)	0.033***	0.001	0.096***	0.001				
railline250	-0.051***	0.001	-0.055***	0.001	-0.050	0.001	-0.047***	0.001
railline250_500	-0.038***	0.001	-0.042***	0.001	-0.037	0.001	-0.036***	0.001
R square	0.829		0.831		0.829		0.830	
N	542,884		543,873		542,884		543,873	

Linear regression model coefficients with standard errors of the estimates in parentheses.
*** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level..

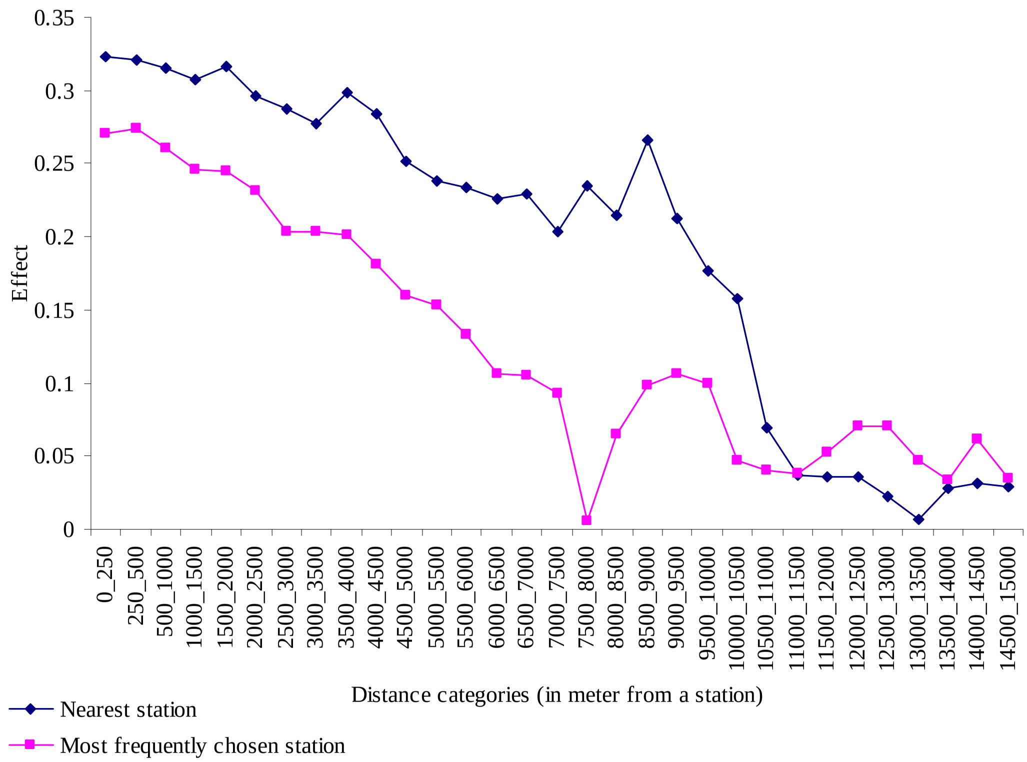


Figure 4.2: Effect of railway station distance on house prices

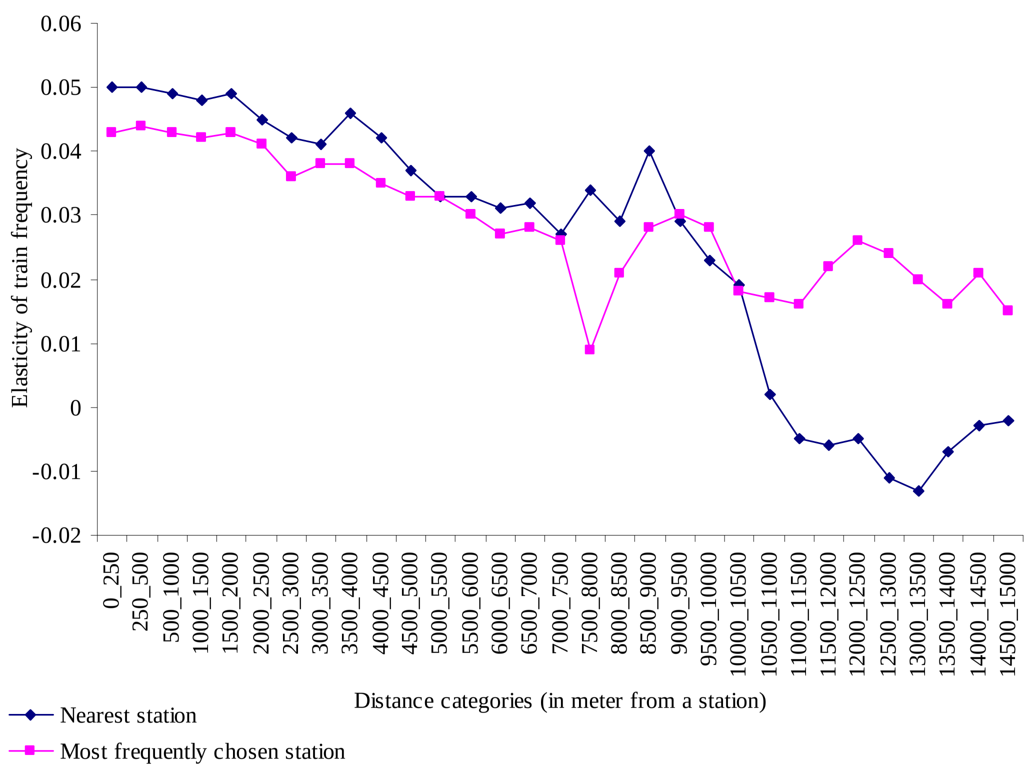


Figure 4.3: Cross-effect of railway station distance and frequency of trains on house prices

Table 4.4: The relative price difference of dwellings at sample distances compared with dwellings located beyond 15 kilometres (based on cross-distance –frequency specification)

	0-250 m		5000-5500 m		10000-10500 m	
Distance Frequency (trains/day)	Nearest station	Mostly chosen station	Nearest station	Mostly chosen station	Nearest station	Mostly chosen station
50	19.6%	16.8%	12.9%	12.9%	7.4%	7.0%
100	23.0%	19.8%	15.2%	15.2%	8.7%	8.3%
200	26.5%	22.8%	17.5%	17.5%	10.1%	9.5%
400	30.0%	25.8%	19.8%	19.8%	11.4%	10.8%
800	33.4%	28.7%	22.1%	22.1%	12.7%	12.0%

To achieve an increase in real estate values along a railway line, there are several strategies. One strategy would be to increase the frequency of service at existing stations, and Table 4.4 (The choice the three distance categories is made solely for comparison purposes) shows the rather modest effects. Another strategy would be to create an extra station. If two stations are located at distances of, say, 10 kilometres, and a new station is built in-between the two, the distance to the nearest station decreases up to a maximum of 5 km. As indicated by Table 4.4, the latter strategy would lead to an increase in the house value of at most 6.7% (19.6%-12.9%) of the dwellings located in the immediate vicinity of the station. With the present model, however, it is not possible to investigate the consequences of adverse effects on travel times as a result of the extra stop. Note that, when we compare the effects of creating an extra station or increasing the frequency of trains, the first mainly affects property values in one location, whereas the latter would be beneficial for all stations at which the train stopped.

4.6 SUMMARY AND CONCLUSION

This chapter has analysed the effect of railway station accessibility on house prices. A cross-sectional hedonic price model is estimated based on Dutch residential house transactions in the years from 1985 to 2001. The model accounts for physical, environmental, temporal and accessibility features of the residential properties. For each of these features, a wide range of variables is included. The main focus of this chapter is, however, to analyse the effect of accessibility provided by railway transport on property values. Most studies in this area only consider the proximity of properties to railway stations. But, this approach is limited because the accessibility of railway stations is more than just proximity to railway stations. In other words, railway stations are not chosen as departure points for reasons of proximity alone. Thus, we need a better approach to address railway accessibility in the analysis. Railway accessibility is a function of the distance and the service levels at the relevant departure railway stations. The choice of a departure railway station is also affected by the levels of rail service, network connectivity, service coverage, and facilities. Thus, it is possible for residential property values to react to an important railway station located farther away than to a less important one located nearby. In this respect, most previous studies have shortcomings in that they neglect the choice process for a departure station in their property value effect analysis by sticking to the nearest railway station. This chapter adds to the literature in this area in two respects. First, we make a distinction between the nearest railway station to the property and the most frequently-chosen station in the postcode area to which the property under consideration belongs. Second, a broader approach for addressing accessibility is applied by taking into account the frequency of train services. The effects of proximity and service levels on property values are analysed. In addition we pay attention to the perpendicular distance to railway lines in order to reflect potential noise and other disturbance effects.

Correcting for a wide range of other determinants of house prices, we find that dwellings very close to a station are, on average, about 25% more expensive than dwellings at a distance of 15 kilometres or more. This percentage ranges between 19% for low-frequency stations and 33% for high-frequency stations (see Table 4.4). A doubling of train frequency leads to an increase of house values of about 2.5%, ranging from 3.5% for houses close to the station to 1.3% for houses further away. Finally, we find a negative effect of distance to railways, probably due to noise effects: within the zone up to 250 metres around a railway line prices are about 5% lower compared with locations further away than 500 metres. As a result of the two distance effects, the price gradient starts to increase as one moves away from a station, followed by a gradual decrease after a distance of about 250 metres.

Our estimations reveal that the distinction between nearest railway station and most frequently-chosen railway station is important. In many cases, the traveller does not choose the closest station. This is an indication that railway station accessibility is a more complex concept than one might think, as it involves competition between railway stations. Further improvement can be done in two areas. First, railway services provided at a railway station are more than just frequency of service. Network connectivity and service coverage in relation to important destination points are an inseparable part of rail services. Thus, to assess railway accessibility a more comprehensive measure that reflects all sorts of rail services provided at a station should be determined. This will be dealt with in the next chapter (Chapter 5). In addition, travellers mostly have a set of railway stations which they use as departure stations to choose from. At the same time, accessing a railway station can be done by different modes of transport. Therefore, the accessibility of a location (a house, etc) to railway transport is explained by a number of factors related to the ease of reaching the railway station in an individual's choice set and the rail and supplementary services provided at the railway stations. The general railway accessibility is therefore an aggregate function of these features over the entire group of railway stations in the choice set, weighted according their importance. Thus, based on both access mode and departure railway station choices, a nested logit model is estimated with the ultimate aim of computing the general railway accessibility at a location. This subject will be covered later in Chapter 6.

APPENDIX 4AI: Transcendental logarithmic formulation

The transcendental logarithmic formulations produce smooth curves, showing the general approximation of effect. We accommodate the distance and frequency of trains in the translog treatment:

$$\begin{aligned}\ln(\mathbf{tranPrice}) = {} & \alpha + \boldsymbol{\beta}'_{\mathbf{HC}} \times \mathbf{HouseChr} + \beta_d \times \ln Rail + \beta_{dSQ} \times (\ln Rail)^2 \\ & + \beta_{freq} \times \ln FreqT + \beta_{freqSQ} \times (\ln FreqT)^2 + \beta_d \times \ln Rail \times \ln FreqT \\ & + \beta_{hw} \times \ln highway + \boldsymbol{\beta}'_{\mathbf{railline}} \times \mathbf{Drailline} + \boldsymbol{\beta}'_{\mathbf{Neighb}} \times \mathbf{Neighb} \\ & + \boldsymbol{\beta}'_{\mathbf{Region}} \times \mathbf{Dregional} + \boldsymbol{\beta}'_{\mathbf{time}} \times \mathbf{Dtime} + \boldsymbol{\varepsilon}.\end{aligned} \tag{4}$$

We also estimate a complete translog formulation, which includes the highway distance to the model as follows:

$$\begin{aligned}\ln(\mathbf{tranPrice}) = {} & \alpha + \boldsymbol{\beta}'_{\mathbf{HC}} \times \mathbf{HouseChr} + \beta_d \times \ln Rail + \beta_{dSQ} \times (\ln Rail)^2 + \beta_{freq} \times \ln(FreqT) \\ & + \beta_{freqSQ} \times (\ln FreqT)^2 + \beta_{hw} \times \ln highway + \beta_{hwSQ} \times (\ln highway)^2 \\ & + \beta_{CrossRailFreq} \times \ln Rail \times \ln FreqT + \beta_{Crossrailhighw} \times \ln Rail \times \ln highw \\ & + \beta_{CrossFreqHighw} \times \ln highw \times \ln FreqT + \boldsymbol{\beta}'_{\mathbf{railline}} \times \mathbf{Drailline} + \boldsymbol{\beta}'_{\mathbf{Neighb}} \times \mathbf{Neighb} \\ & + \boldsymbol{\beta}'_{\mathbf{Region}} \times \mathbf{Dregional} + \boldsymbol{\beta}'_{\mathbf{time}} \times \mathbf{Dtime} + \boldsymbol{\varepsilon}.\end{aligned} \tag{5}$$

In Equations 4 and 5, ‘$Rail$’ is the distance to the railway station in its continuous form and ‘$highway$’ is the distance to the highway entry/exit point; the remaining variables are defined in Section 4.4 above.

Table 4AI. 1Estimation of railway station effect on house values: transcendental logarithmic formulation

	Nearest station	**Most frequently Chosen station**	**Nearest station**	**Most frequently Chosen station**
(Constant)	8.863*** (0.044)	8.422*** (0.036)	9.673*** (0.070)	9.391*** (0.055)
Log (railway station dist)	0.198*** (0.007)	0.203*** (0.005)	0.233*** (0.009)	0.317*** (0.007)
Log (railway station dist) square	-0.019*** (0.000)	-0.024*** (0.000)	-0.018*** (0.000)	-0.025*** (0.000)
Log (frequency)	0.037*** (0.011)	0.152*** (0.010)	-0.332*** (0.013)	-0.220*** (0.011)
Log (frequency) square	-0.008*** (0.001)	-0.019*** (0.001)	0.000 (0.001)	-0.009*** (0.001)
Log (highway dist)	0.014*** (0.001)	0.024*** (0.001)	0.014 (0.009)	-0.073*** (0.009)
Log (highway dist) square			-0.012*** (0.000)	-0.005*** (0.000)
log (railway station dist)* log (frequency)	0.011*** (0.001)	0.020*** (0.001)	0.006*** (0.001)	0.010*** (0.001)
Log (railway station dist)*log (highway dist)			-0.004*** (0.001)	-0.007*** (0.001)
Log (frequency)*log (highway dist)			0.042*** (0.001)	0.043*** (0.001)
railline250	-0.044*** (0.001)	-0.055*** (0.001)	-0.049*** (0.001)	-0.057*** (0.001)
railline250_500	-0.037*** (0.001)	-0.047*** (0.001)	-0.041*** (0.001)	-0.047*** (0.001)
Log surface area	0.208*** (0.001)	0.213*** (0.001)	0.209*** (0.001)	0.213*** (0.001)
Building age	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
Log (number of rooms)	0.300*** (0.001)	0.298*** (0.001)	0.301*** (0.001)	0.299*** (0.001)
Number of bathrooms	0.090*** (0.001)	0.089*** (0.001)	0.090*** (0.001)	0.089*** (0.001)
Presence of gas heater	-0.147*** (0.001)	-0.147*** (0.001)	-0.145*** (0.001)	-0.145*** (0.001)
Presence of open fireplace	0.065*** (0.001)	0.063*** (0.001)	0.065*** (0.001)	0.063*** (0.001)
Presence of monument	0.305*** (0.004)	0.299*** (0.004)	0.286*** (0.004)	0.284*** (0.004)
Presence of garage	0.106*** (0.001)	0.108*** (0.001)	0.106*** (0.001)	0.107*** (0.001)
Presence of garden	0.024*** (0.001)	0.024*** (0.001)	0.023*** (0.001)	0.024*** (0.001)
R-square	**.827**	**.830**	**.828**	**.831**
Number of observations	**542884**	**542884**	**542884**	**542884**

Linear regression model coefficients with standard errors of the estimates in parentheses.
*** significant at the 1% level.
** significant at the 5% level.
* significant at the 10% level..

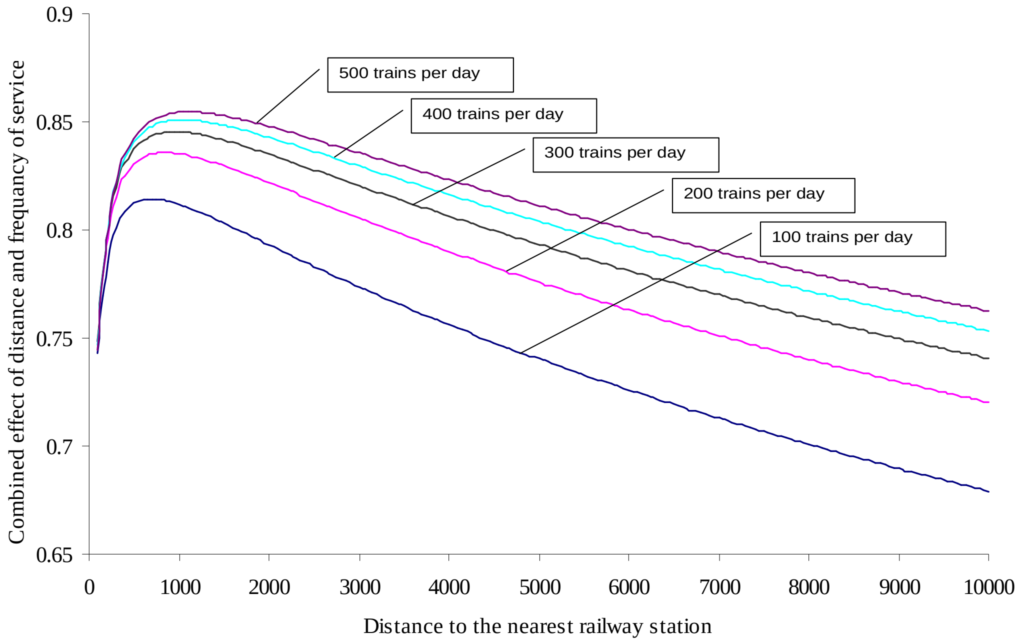


Figure 4AI. 1: Effect of distance and frequency of trains based on the nearest station.

The use of the translog function does not give a detailed treatment of the effect of distance; this can be done in a better way by the stepwise distance functions reported in Table 4.3 in the main text. However, the translog model is better in dealing with the effect of frequency, in particular the extent to which frequency effects are different for houses close to stations and houses further away. On the Y-axis of Figure 4AI.1, we have value of the log price determined as the combined effect of distance to the railway station and frequency in the translog formulation given above. Because of the multiplicative nature of the specified model, the monetary or percentage effect of distance and frequency of trains at the stations can not be inferred from the graphs. However, the graphs reveal the general pattern of distance and frequency of train effect. Figure 4AI.1 is based on the effect of the nearest station given in column 1 of Table 4AI.1. On the X-axis we have distance to the station (in this case to the nearest). The curves represent the different levels of frequency of trains at the nearest stations. The lower curve corresponds to a frequency level of 100 trains per day, whereas the upper curve corresponds to a frequency of 500 trains per day. The frequency interval between the curves is fixed to 100 trains per day to facilitate comparison concerning the effect of additional trains. Figure 4AI.1 shows that not only does low frequency lead to a lower house price, but also that, for low frequencies, the distance decay is faster. As we move from the lower-level of frequency to the highest level of frequency we observe a diminishing contribution of frequency to the log of house price of a given location. On the other hand, the increase in the frequency levels has an increasing effect on the log of house price with increase in distance. A doubling of frequency from 100 to 200 trains per day has an effect of about 2.9% on the log of house price at a distance of 1000 metres, whereas this effect is about 4.8% at about 5 km and 6% at 10 km. In addition, the general structure of the curves indicates that the houses located immediately adjacent to the stations sell at lower prices than houses located some few hundreds of metres from the station.

The graphs also enable us to compare the effect of distance and frequency of trains at the station. Consider the log price value corresponding to the 100 trains per day frequency at a distance of 1 kilometre from the station. This is found to be equivalent to a value at 2.6 kilometres and a frequency level of 200 trains per day. Thus, according to market valuations, a doubling of frequency has a value that is about equal to a reduction of distance of about 1600 metres for this case.

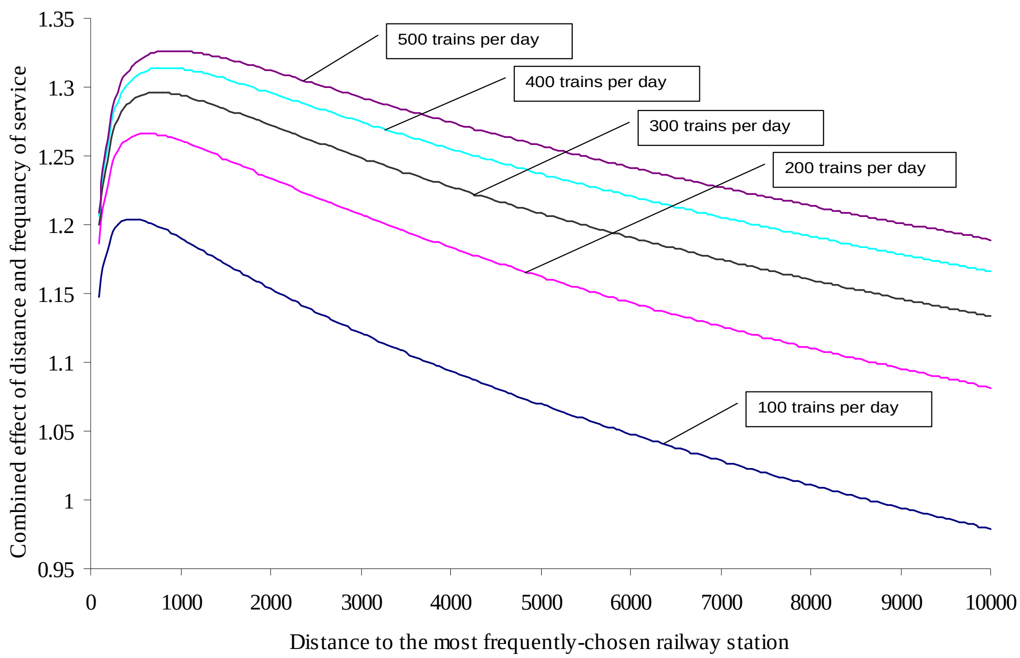


Figure 4AI. 2: Effect of distance and frequency of trains based on the most frequently-chosen station.

Figure 4AI.2 above is based on column 2 of Table 4AI.1. It shows the effect of distance and frequency of trains at the most frequently-chosen station on house prices. The general structure of the curves remains the same as the curves based on the nearest station. The main difference between the two is shown by the value of the total effect of distance and frequency of trains on house prices. The most frequently-chosen station results in a higher total effect on house prices compared with an effect produced by the nearest station.

Chapter 5

5 A Measure of Railway station's Service Quality

5.1 INTRODUCTION

Transport infrastructure is broadly defined as transport related capital that provides public services (Rietveld and Bruinsma 1998). It has drawn attention in many scientific researches from different angles. One stream of the research is devoted to the analysis of the impact of transport infrastructure. The focus of infrastructure impact can take different form: a broad macro perspective such as the impact of transport infrastructure on the economy of a certain geographical area, employment etc. On the other hand, the analysis of infrastructure impact can assume a micro perspective. One such example is the analysis on the impact of transport infrastructure on property value. The impact of transport infrastructure on property values come from the resulting improvements in the accessibility level. Accessibility is generally defined as the potential of opportunity for interaction. Mostly, accessibility is assessed in reference to nodes in a transport network. However, there are several operational definitions which are adopted in the literature. In this chapter we adopt the operational definition of accessibility as the net aggregate weighted travel services provided by a transport node. Talking in terms of transport infrastructure can still be broad in that transport infrastructure encompasses different modes of transport. This thesis is devoted to analyse the impact of railway infrastructure. In this chapter we try to quantify the railway related accessibility level provided by railway stations in the Dutch railway network.

When we speak about the impact of transport infrastructure on property values, we have to be clear about the source of the impact. It is not the investment in itself that affects property values, but the transportation-benefit services supplied as a result of the investment in the transportation infrastructure. Thus, in this regard, quantifying transportation infrastructure means quantifying the transportation benefit due the transportation infrastructure. Railway stations can be treated as the outlets where railway services from railway infrastructure are delivered. Thus, quantifying the railway service provided at the station enables us to measure the benefit of the rail infrastructure to the travellers at that point in space.

Railway stations differ from each other in several respects. In the literature, a typical distinction between railway stations is made with respect to the type of railway station. Four types of rail transit stations can be identified: commuter railway stations; heavy railway stations; light railway stations; and bus rapid transit stations (BRT) (see Chapter 2). Even with such distinctions, we observe heterogeneity among stations of the same type. For instance, in the Netherlands there are four types of commuter railway stations: namely, the all-station 'stop-train' rail services; semi-fast also called 'express' rail services, which call at main and medium sized cities; intercity rail services that only call at main cities; and international trains that only stop at a very limited number of stations. Moreover, it is also known that the rail service levels of stations of the same type can differ. Thus, there is a need for a refined method of distinguishing the features of railway stations for a proper analysis regarding railway accessibility and departure station choice. The first step is to identify rail service features that have railway accessibility implications. Generally speaking, the services provided by the railways are of two types. The first type of services relate to pure rail services provided at the station. The second type relates to the supplementary services made available for railway travellers at the station. This type includes services such as the availability of parking spaces, the park-and-ride possibility, bike stands and storage facilities. These services can be provided by the Railway Company or local authority. However, the fact that they are provided in relation to the railway station makes them part of the services provided at the stations. The focus of this chapter will be on the pure rail services provided at the stations.

In the context of pure rail services, the service provided at a station can be assumed to relate to three aspects which have implications for the total travel time. First, it relates to how quickly travellers can get service. In other words, this means the average time that travellers have to wait before catching a train. This feature is determined by the frequency of trains leaving the station per a period of time. A shorter waiting time implies the importance of the railway station as a departure point. Second, it relates to how well the station in consideration is connected to other stations in the network. This indicates the level of service coverage provided by the railway station. In addition, the service level can be related to the level of (network) connectivity to other stations in the network. The number of direct connections from the station is a good indicator of the network connectivity of a station. However, some stations can only be accessed though a transfer station. Thus, the time lost in changing trains is a good indicator of how well the station is connected to other stations in the network. Third, the rail service provided at a station is related to the relative position of the station in the network. This feature has a direct relation with the distance between stations and the speeds at which the trains operate. The in-vehicle travel time is an important determinant of this feature. Furthermore, being close to important destination stations in the network increases the attractiveness of a station as a departure point.

The aim of this chapter is to develop a comprehensive rail service quality index (RSQI) of a station in a network. This RSQI will be used in a subsequent choice analysis for departure stations and real estate price analyses (see Chapters 6, 7 and 8). Section 5.2 discusses the model applied in this chapter. Section 5.3 discusses the data used for the estimation. This will be followed by the estimation and discussion of the results (see Section 5.4). The chapter ends with a conclusion.

5.2 RAIL SERVICE QUALITY INDEX

Railway stations differ from each other in the rail services they offer to passengers. In many empirical applications it has been noted that there is a need to distinguish between stations on the basis of the service levels. In our meta-analysis discussion (see Chapter 2), we noted that the intensity of the effect exerted by railway stations differs from one type to another. On the other hand, the findings of Chapter 4 indicate that the frequency of rail service as an indicator of rail service provided in railway stations is an important factor in determining real estate prices together with the proximity to railway stations. In addition to the frequency of rail service, the data set includes features such as the number of destinations having a direct connection with the station under consideration, and the intercity status of the station. However, the usefulness of these factors in accounting for rail service is limited because they do not take the location of the station in relation to other (important) stations in the network into account. Thus, in addition to the factors mentioned above, the importance of the other stations and distance from the other stations are important factors in determining the service quality of a station. The need for a comprehensive rail service quality indicator for each railway station leads us to an estimation exercise based on the underlying railway trip data. We call this index the Rail Service Quality Index $(RSQI)$. Below we discuss the RSQI of a station both from a departure and destination view points.

5.2.1 Departure station

The importance of a station as a departure point is valued by the access it provides to a wide range of destinations. At the same time the importance of destinations can differ considerably. The importance of a destination station can be measured by the trip-attracting capacity of the station. Stations which have higher trip attracting capacity are presumed to be important destinations. Therefore, if a railway station enjoys good connections to stations which have high trip-attraction capacity, it is said to have good rail service quality. In addition to having a train service directed to these destinations, connectedness may also imply lower generalized journey time (for an explanation of this term see text below Equation 1) and a lower journey time to distance ratio. Thus, the level of the pure RSQI of a station as a departure point is a function of the importance of the destination stations, the generalized journey time it takes to reach the stations and the ratio of generalized journey time to distance. The importance of a destination station can be explained by the size of the station as a destination point. The overall rail service quality indicator of a departure station is therefore an aggregate sum of the function over all destination railway stations:

$$RSQIdeparture_i = \sum_j f(D_j, GJT_{ij}, GJT_{ij} / d_{ij}), \tag{1}$$

where, D_j is the total number of trips attracted by a destination station j; GJT_{ij} is the generalized journey time between stations i and j: generalized journey time is a measure of the time needed to travel between stations. It includes the average waiting time, in-vehicle time, transfer time and some penalty for the number of transfers. The generalized journey time measure encompasses several station-distinctive features of stations. For instance, the ‘frequency of trains leaving the station per period of time’ is reflected in the average waiting time component. The distinction of railway stations as ‘intercity’, ‘semi-fast’ and ‘stop-train’ is expected to be reflected in the transfer time and number of transfer penalties. Intercity train stations provide more direct services. This leads to less in-vehicle and transfer time, and thus, less generalized travel time than semi-fast and stop train stations. In addition, the ‘connection time’ and ‘penalty for the number of connections’ shows the level of direct connection a station has with other stations. Generalized journey time is expected to have a negative effect on the general rail service quality of a railway station. The shorter the time it takes to reach the destination stations from the departure station concerned the higher is the rail service quality of the railway station. GJT_{ij} / d_{ij} is the ratio of generalized journey time to the distance between stations i and j. Distance in this function is given by the Euclidian distance between the two stations. The generalized journey time to distance ratio in the function is used to control for the effect of other modes of transport on the general attractiveness of railway transport. A high value of the ratio of generalized journey time to distance implies that the train trip involves a larger detour to reach the destination station. This opens the possibility of substituting the train by other modes of transport. Thus, it has a negative effect on the attractiveness or the general rail service quality of a station.

5.2.2 Destination station

The quality of a station as a destination station is determined by its accessibility to trips ending at the station. As distinguished from the departure-station quality of a station discussed above, the importance of a station as a destination station is affected by the size of the origin stations, generalized journey time, and the ratio of generalized journey time to distance:

$$RSQIdestination_j = \sum_i f(O_i, GJT_{ij}, GJT_{ij} / d_{ij}), \qquad (2)$$

where, O_i is the total number of trips originateing in station i. This is an indicator of the importance of the origin station to which the destination station in consideration is connected. The remaining variables are as explained above.

5.3 SPATIAL INTERACTION MODELS

Spatial interaction models are designed to model the trip distribution between stations. They aim to explain the factors that promote or discourage flow nodes. In addition, they can be applied to predict the flows between nodes for a given change in the settings of the factors that affect flow distribution. In terms of explanation applicability, spatial interaction models are grouped into three types: 1) models that provide information only on destination station features. These models are generally known as production-constrained models; 2) models that provide information only on origin station features. These are called attraction constrained models; and 3) models that provide information on both origin and destination features. These models are termed unconstrained models. Yet a fourth type of interaction models exists. It is mostly used for flow prediction purposes rather than explanation. Models of this type are constrained at both from the origin and the destination nodes. These models are known as the doubly-constrained interaction models (Fotheringham and O'Kelly 1989; and Ortuzar and Willumsen 2001). They utilize the trip production capacity of the origin and the trip attraction capacity of the destination station as exogenous variables. It is believed that these models give a higher level of prediction accuracy than the other three types of models. For a detailed discussion on the interaction models, refer to Fotheringham and O'Kelly (1989). In this chapter we use the doubly-constrained spatial interaction model for estimating the parameters that we use in determining the RSQI of a railway station.

Doubly-constrained model

As the name implies, the model is constrained at both the origin and the destination stations. The constraint pertains to the production capacity of an origin station and the attraction capacity of a destination station. These capacities are constrained to be equal to the sum of all trips originating at the departure station and those ending at the destination station, respectively. It has been explained earlier that, apart from the nature of the destination or origin, station flow between stations is affected by the generalized journey time and the ratio of generalized journey time to distance. The general form of the doubly-constrained model used to model spatial interaction between stations is given as follows:

$$T_{ij} = A_i O_i B_j D_j \, f(GJT_{ij}) \, f(\frac{GJT_{ij}}{d_{ij}}) \exp(\xi_{ij}); \tag{3}$$

$$O_i = \sum_j T_{ij}; \tag{4}$$

$$D_j = \sum_i T_{ij}. \tag{5}$$

where T_{ij} is the number of trips between the stations origin station i and destination station j; A_i and B_j are the balancing factors which ensure that the constraints on origins and destinations (given by Equations 4 and 5) are met; O_i is the total number of trips originating in station i; D_j is the total number of trips attracted by a destination station j; $f(GJT_{ij})$ is a function of the generalized journey time between stations i and j; and $f(GJT_{ij} / d_{ij})$ is a function of the generalized journey time and the distance ratio between stations i and j; and, lastly ξ_{ij} is the error component of the model which follows an independently and identically normal distribution.

In order to estimate the model, it is necessary to select a form for the functions of the generalized journey time and ratio of the generalized journey time to distance. The specifications are given in Equations 6 and 7.

$$f(GJT_{ij}) = \sum_{c=1}^{C} \beta_c DGJT_c^{\ ij}, \tag{6}$$

This is a stepwise function of the generalized journey time. $DGJT_c^{\ ij}$ is a dummy variable which is equal to 1 if GJT_{ij} falls in the generalized journey time category c, and 0 otherwise. β_c is the coefficient for generalized journey time category c:

$$f\left(\frac{GJT_{ij}}{d_{ij}}\right) = \left(\frac{GJT_{ij}}{d_{ij}}\right)^{\gamma} \tag{7}$$

where, γ is the power coefficient the ratio of generalized journey time and distance. Thus, the doubly-constrained gravity model that we estimate is given by:

$$T_{ij} = A_i O_i B_j D_j \left(\sum_{c=1}^{C} \beta_c DGJT_c^{\ ij}\right)\left(\frac{GJT_{ij}}{d_{ij}}\right)^{\gamma} \exp(\xi_{ij}). \tag{8}$$

This equation can be linearized by taking the natural logarithm of both sides:

$$\ln\left(\frac{T_{ij}}{O_i D_j}\right) = \ln A_i + \ln B_j + \ln\left(\sum_{c=1}^{C} \beta_c DGJT_c^{\ ij}\right) + \gamma \ln\left(\frac{GJT_{ij}}{d_{ij}}\right) + \xi_{ij}. \tag{9}$$

The coefficient of the generalized journey time categories, the ratio of generalized journey time, and the balancing factors will be estimated from the above equation. Thus, in the estimation the logs of the balancing factors in the equation represent the coefficients to be estimated. This requires that the logs of the balancing factors are multiplied by the dummy variable for the corresponding station. Therefore, the equation to be estimated is given as follows:

$$\ln\left(\frac{T_{ij}}{O_i D_j}\right) = \sum_{\tilde{i}=1}^{N} \ln A_{\tilde{i}} S_{\tilde{i}} + \sum_{\tilde{j}=1}^{N} \ln B_{\tilde{j}} S_{\tilde{j}} + \ln\left(\sum_{c=1}^{C} \beta_c DGJT_c^{\ ij}\right) + \gamma \ln\left(\frac{GJT_{ij}}{d_{ij}}\right) + \xi_{ij}. \tag{10}$$

where, N is the number of railway stations in the railway network; and $S_{\tilde{i}}$ and $S_{\tilde{j}}$ are dummy variables for departure station $\tilde{i}$ and destination station $\tilde{j}$. They assume the value 1 when $i = \tilde{i}$ and $j = \tilde{j}$, respectively, and 0 otherwise. Given the assumption on the error components above, Equation 10 can be estimated using ordinary least squares (OLS). The estimated coefficients are then used in determining the RSQIs for each station. As pointed out earlier, the RSQI of a station can, however, be viewed from two angles: whether the station is treated as a departure station or as destination station. We make this distinction at this point because the two indices have different implications for different type of real estate analysis. For instance, for a residential-property value analysis, the departure-station perspective of the service quality is relevant for the analysis. On the other hand, a commercial-property value analysis requires the treatment of the rail service quality of a station as a destination station. From a departure station setting the index is determined by the generalized journey time, the size of the destination station, given by the trips attracted by the destination station, and the generalized journey time and distance ratio. On the other hand, from a destination-station setting, the index is determined by replacing the size of the origin station in place of the size of the destination station mentioned above. An aggregation over all the destination stations from stations j and origin stations i gives the value of both indices, respectively. The RSQIs of a station as a departure point and a destination point are specified in Equations 11 and 12, respectively:

$$SQIdeparture_i = \sum_j \hat{B}_j D_j \hat{f}(GJT_{ij}) \hat{f}(\frac{GJT_{ij}}{d_{ij}}) = \sum_j \hat{B}_j D_j \left(\sum_{c=1}^{C} \hat{\beta}_c DGJT_c^{\ ij}\right)\left(\frac{GJT_{ij}}{d_{ij}}\right)^{\hat{\gamma}}; \tag{11}$$

$$SQIdestination_j = \sum_i \hat{A}_i O_i \hat{f}(GJT_{ij}) \hat{f}(\frac{GJT_{ij}}{d_{ij}}) = \sum_i \hat{A}_i O_i \left(\sum_{c=1}^{C} \hat{\beta}_c DGJT_c^{\ ij}\right)\left(\frac{GJT_{ij}}{d_{ij}}\right)^{\hat{\gamma}}. \tag{12}$$

5.4 ESTIMATION AND DISCUSSION

The estimation of the doubly-constrained model given by Equation 10 is based on train trips from 365 departure railway stations to 365 destination train stations. These stations are all the stations in the Dutch domestic railway network. International destinations are not included. This may cause an understatement of the railway service quality for some stations which have important international connections. However, the model is flexible enough to accommodate all stations accessed from a particular station. The data used in our estimation are acquired from the Dutch Railway company (Nederlandse Spoorwegen - NS). The data set includes the number of trips, generalized journey time, and distance between each pair of stations. Trips originating and ending at a station are determined by the aggregation of the trips over all destination and departure stations, respectively. The descriptive statistics of variables used in our estimation are given in Table 5.1 below

Table 5.1: Descriptive statistics of railway station database

	Minimum	Maximum	Mean	Std. Deviation
Generalized journey time (in minutes)	6	454	178	72
Distance between stations (kilometres)	0.68	312.90	107.70	58.64
Time to distance ratio	0.75	73.32	1.98	1.27
Production capacity (passengers)	4,495	7,977,940	437,483	853,611
Attraction capacity (passengers)	1,004	15,554,143	438,362	1,362,395

In our estimation, the generalized journey time variable is divided into 46 categories of 10-minute intervals. The categories assume a dummy value of 1 if the generalized journey time of trips between any pair of stations falls within a range corresponding to the category, and 0 otherwise. The last category is taken as a reference group. During the analysis it was necessary to make some computational adjustments. This is because for some pairs of stations, there were no trips. Taking the logarithm of these values leads to the exclusion of these entries from the estimation. To avoid this problem, a small value had to be added to find a positive value for the number of trips between the pairs of stations. Sen and Smith (1995) have proved that the optimal value that can be added is ½ a trip. Following that conclusion, our final estimation of the parameters is based on the actual trips plus ½ a trip. The estimation result of the doubly-constrained interaction model (10) is given in Table 5.2. The table only gives the coefficients of time categories and the ratio of time to distance. The balancing factors are not reported here. The coefficients of the generalized journey time categories given in the table represent the natural logarithms of the actual coefficients (see Equation 10). All coefficients are significant and with the expected sign.

Table 5.2: Estimation result of the doubly-constrained interaction model

Variable	Coefficient	Std. Error	t	Sig.
GJT0_10	8.642	0.411	21.027	0.000
GJT10_20	9.261	0.373	24.837	0.000
GJT20_30	9.351	0.372	25.160	0.000
GJT30_40	8.736	0.371	23.525	0.000
GJT40_50	8.114	0.371	21.859	0.000
GJT50_60	7.561	0.371	20.378	0.000
GJT60_70	7.025	0.371	18.943	0.000
GJT70_80	6.414	0.371	17.302	0.000
GJT80_90	5.917	0.371	15.965	0.000
GJT90_100	5.482	0.371	14.796	0.000
GJT100_110	5.066	0.370	13.676	0.000
GJT110_120	4.684	0.370	12.644	0.000
GJT120_130	4.424	0.370	11.944	0.000
GJT130_140	4.137	0.370	11.170	0.000
GJT140_150	3.913	0.370	10.567	0.000
GJT150_160	3.742	0.370	10.106	0.000
GJT160_170	3.550	0.370	9.588	0.000
GJT170_180	3.391	0.370	9.157	0.000
GJT180_190	3.214	0.370	8.682	0.000
GJT190_200	3.083	0.370	8.328	0.000
GJT200_210	2.850	0.370	7.697	0.000
GJT210_220	2.739	0.370	7.397	0.000
GJT220_230	2.552	0.370	6.893	0.000
GJT230_240	2.386	0.370	6.444	0.000
GJT240_250	2.208	0.370	5.964	0.000
GJT250_260	2.061	0.370	5.566	0.000
GJT260_270	1.930	0.370	5.212	0.000
GJT270_280	1.829	0.370	4.939	0.000
GJT280_290	1.623	0.370	4.382	0.000
GJT290_300	1.539	0.370	4.155	0.000
GJT300_310	1.369	0.371	3.693	0.000
GJT310_320	1.280	0.371	3.451	0.001
GJT320_330	1.094	0.371	2.946	0.003
GJT330_340	0.995	0.372	2.675	0.007
GJT340_350	0.913	0.372	2.452	0.014
GJT350_360	0.842	0.372	2.261	0.024
GJT360_370	0.758	0.372	2.035	0.042
GJT370_380	0.718	0.373	1.927	0.054
GJT380_390	0.823	0.374	2.203	0.028
GJT390_400	0.831	0.375	2.214	0.027
GJT400_410	0.681	0.377	1.805	0.071
GJT410_420	0.664	0.388	1.712	0.087
GJT420_430	0.730	0.397	1.841	0.066
GJT430_440	0.493	0.422	1.168	0.243
Log(GJT/dist)	-0.399	0.011	-37.449	0.000

Mapping the value of the coefficient for the time categories gives an insight into the effect of time on trips between stations. Naturally, one would expect the number of trips between any pair of stations to decline as the travel time between the stations increases. However, as we can see from Figure 5.1, the graph is an increasing function of travel time for the initial stages. For trip durations of up to 30 minutes, train trips are increasing with time. This indicates that for shorter trips the train encounters competition from other modes. Apparently, the competition effect presented by the generalized journey time to distance ratio does not completely capture the competition phenomena. A possible explanation is that, for shorter trips, people tend to use other modes such as bicycle and public transport rather than the train, even if the train generally involves shorter journey time. Trips between points within cities are generally expected to be accommodated by walking, biking, or public transport because of the flexibility they offer for a multi-purpose trip. As the trip duration increases, train trips are expected to take over. For trip duration of over 30 minutes, the competition effect from other modes more or less disappears, and the real negative effect of time operates. The graph shows a smooth decline in the train trips as the trip time increases.

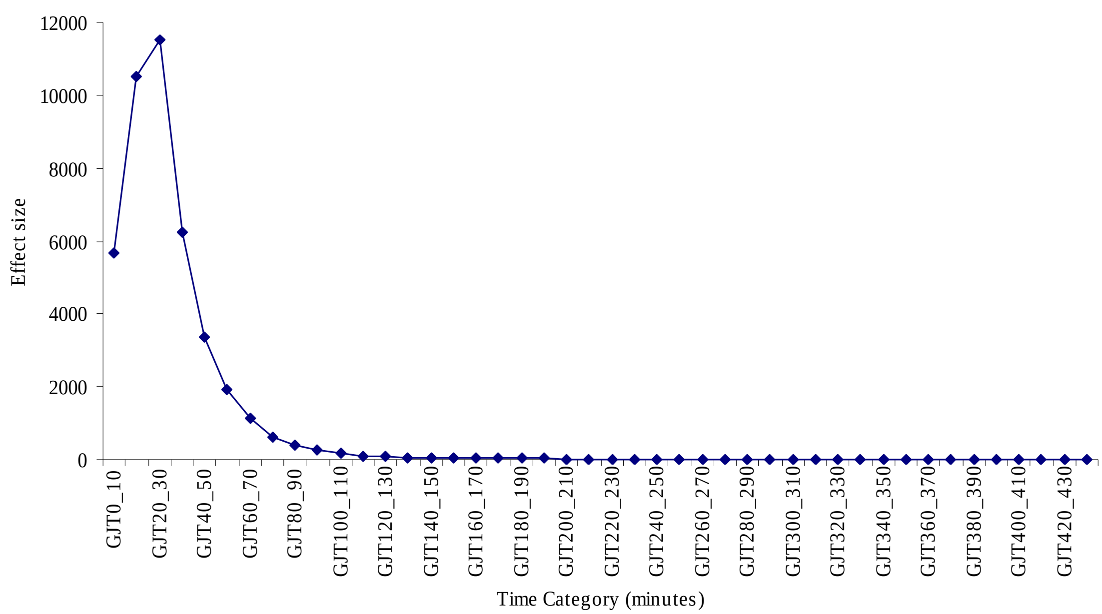


Figure 5.1: Effect of trip duration on train trips

In addition, we can see the effect of the generalized journey time to distance ratio on train trips. Normally, the distance between two stations is fixed. If the train trip involves detours, this implies the generalized travel time increases, and thus the ratio. Figure 5.2 below demonstrates the effect of increasing the generalized journey time to distance ratio on train trips. As the generalized journey time to distance ratio increases, the number of train trips declines. Both competition and travel time effects play a role in the decline in the level of train trips. First, the fact that the train trip involves longer detours makes other modes of transport preferable. For shorter distances, bike and public transport will be preferable. On longer distances, the car option becomes preferable. Second, if the distance is long enough and the train trip involves still further detours, it is generally expected to result in trip losses.

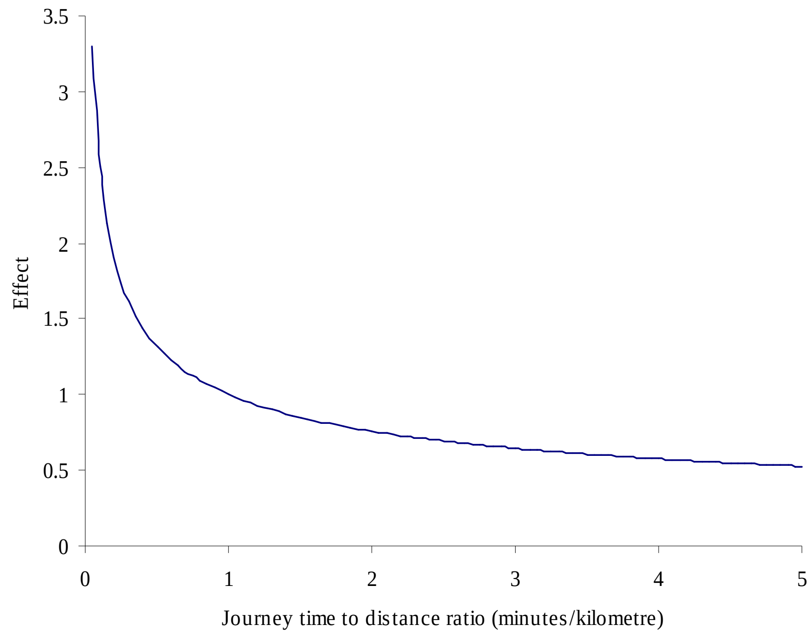


Figure 5.2: Effect of generalized journey time to distance ratio on train trips.

Effect of change in the components of generalized journey time

Using the parameter estimates of the doubly-constrained spatial interaction model, we determine the service RSQI of each railway station in the Dutch railway network. The descriptive summaries of the index for railway stations are given in Table 5.3. The RSQIs for individual railway stations in the Dutch national railway network are given in Table 5AI.1 in the Appendix. The difference in the RSQI of a railway station with the highest value and a station with the lowest value is a factor of about 67 for the departure station under consideration and about 50 for the mean destination station. The effect on RSQI of a doubling of frequency of service to and from a station is considered. In the case of a departure station, a doubling of frequency of service in the network leads to an increase of the RSQI by 0.18. This is 0.14 for the destination station under consideration. In relative terms, these are increases of about 40%. This shows that the model prediction based on the estimation result can result in counter-intuitive results for trips that already take a short time.

Table 5.3: Descriptive statistics of RSQI of a railway station

	N	Minimum	Maximum	Mean	Std. Deviation
RSQI_departure	365	.03	2.00	.44	.33
RSQI_destination	365	.03	1.46	.34	.24

5.5 CONCLUSION

The method of measuring the rail service quality provided by a railway station in a given network that was discussed in this chapter provides a flexible and comprehensive approach. It is comprehensive because it incorporates several features that have rail service features. It is expected to upgrade the quality of measuring railway accessibility in empirical research.

In subsequent chapters, we will use the outcome of this chapter as an input for our analysis. The rail service quality index (RSQI) determined in this chapter is used in the choice analysis for the departure station and access mode discussed in the next chapter. Furthermore, the RSQI will be used to analyse the impact of railway accessibility on the rent levels of office space discussed in Chapter 8.

APPENDIX 5AI: Rail service quality indices (RSQI) of railway station in the Netherlands

Table 5AI. 1: Service quality indices of railway station in the Dutch national railway network

Station name	RSQIdept	RSQIdest
Utrecht Centraal	2.001	1.464
Duivendrecht	1.832	1.269
Leiden Centraal	1.818	1.285
Den Haag HS	1.501	1.118
Schiphol	1.497	1.047
Gouda	1.458	0.884
Haarlem	1.392	0.948
Amsterdam Centraal	1.381	1.058
Amersfoort	1.377	0.915
Weesp	1.351	0.944
s’ Hertogenbosch	1.332	0.953
Amsterdam Sloterdijk	1.286	1.058
Rotterdam Centraal	1.255	1.038
Delft	1.227	0.867
Dordrecht	1.216	0.959
Woerden	1.172	0.862
Heemstede-Aerdenhout	1.144	0.753
Den Haag Centraal	1.144	0.946
Amsterdam Amstel	1.126	0.901
Hoofddorp	1.073	0.746
Naarden-Bussum	1.047	0.740
Hilversum	1.034	0.742
Amsterdam Lelylaan	1.000	0.653
Schiedam Centrum	0.995	0.768
Rijswijk	0.994	0.731
Zaandam	0.992	0.701
Arnhem	0.957	0.807
Amsterdam Zuid WTC	0.948	0.742
Rotterdam Alexander	0.906	0.683
De Vink	0.906	0.710
Diemen Zuid	0.890	0.683
Geldermalsen	0.886	0.616
Delft Zuid	0.876	0.677
Amsterdam Muiderpoort	0.870	0.609
Eindhoven	0.858	0.687
Ede-Wageningen	0.855	0.602
Voorschoten	0.852	0.620
Tilburg	0.847	0.657
Driebergen-Zeist	0.844	0.651
Abcoude	0.842	0.695
Breukelen	0.837	0.679
Den Haag Moerwijk	0.837	0.638
Den Haag Mariahoeve	0.828	0.629
Culemborg	0.826	0.538
Koog Bloemwijk	0.821	0.527
Amsterdam Bijlmer	0.815	0.575
Hilversum Sportpark	0.815	0.577

Station name	RSQIdept	RSQIdest
Koog-Zaandijk	0.802	0.523
Breda	0.796	0.643
Haarlem Spaarnwoude	0.796	0.534
Rotterdam Lombardijen	0.793	0.629
Oss	0.789	0.629
Nieuw Vennep	0.787	0.574
Utrecht Overvecht	0.785	0.608
Amsterdam RAI	0.784	0.605
Wormerveer	0.782	0.494
Boxtel	0.779	0.591
Den Haag Laan van NOI	0.761	0.677
Rotterdam Blaak	0.761	0.642
Houten	0.749	0.538
Zwijndrecht	0.746	0.568
Maarssen	0.745	0.495
Sittard	0.745	0.594
Zoetermeer	0.738	0.546
Barendrecht	0.725	0.538
Hollandsche Rading	0.716	0.478
Almere Centrum	0.704	0.498
Uitgeest	0.702	0.488
Rotterdam Zuid	0.697	0.559
Almere Muziekwijk	0.691	0.494
Nijmegen	0.683	0.585
Assen	0.661	0.516
Elst	0.651	0.533
Voorburg	0.648	0.515
Den Dolder	0.644	0.446
Bunnik	0.639	0.455
Nijkerk	0.639	0.422
Weert	0.634	0.443
Bussum Zuid	0.633	0.441
Tilburg West	0.631	0.529
Almelo	0.631	0.424
Amersfoort Schothorst	0.631	0.475
Baarn	0.628	0.461
Diemen	0.627	0.420
Hillegom	0.625	0.492
Schiedam Nieuwland	0.623	0.485
Zwolle	0.621	0.547
Bilthoven	0.617	0.433
Deventer	0.613	0.488
Gouda Goverwelle	0.606	0.423
Capelle Schollevaar	0.606	0.410
Gilze-Rijen	0.605	0.500
Hilversum Noord	0.597	0.424
Vleuten	0.595	0.477

Table 5AI. 2: Service quality indices of railway station in the Dutch national railway network (Continued)

Station name	RSQIdeprt	RSQIdest
Vlaardingen Oost	0.594	0.463
Almere Buiten	0.591	0.440
Voorhout	0.586	0.430
Utrecht Lunetten	0.584	0.408
Zutphen	0.584	0.537
Zoetermeer Oost	0.574	0.367
Roermond	0.570	0.479
Leidschendam-Voorburg	0.565	0.465
Bodegraven	0.565	0.417
Bloemendaal	0.563	0.372
Dieren	0.561	0.521
Vlaardingen West	0.560	0.444
Maassluis	0.560	0.434
Alphen aan den Rijn	0.558	0.456
Vlaardingen Centrum	0.554	0.455
Roosendaal	0.548	0.445
Voorburg 't Loo	0.548	0.387
Santpoort Zuid	0.547	0.362
Krommenie-Assendelft	0.546	0.379
Best	0.545	0.399
Castricum	0.529	0.402
Alkmaar	0.527	0.415
Zoetermeer Voorweg	0.517	0.348
Putten	0.514	0.325
Eindhoven Beukenlaan	0.513	0.350
Zaandam Kogerveld	0.511	0.356
Zoetermeer Centrum West	0.507	0.346
Beverwijk	0.505	0.429
Nieuwerkerk a/d IJssel	0.501	0.366
Zuidhorn	0.499	0.375
Arnhem Velperpoort	0.495	0.393
Helmond	0.492	0.355
Beek-Elsloo	0.485	0.389
Ravenstein	0.483	0.406
Hengelo	0.482	0.405
Haren	0.481	0.380
Almere Parkwijk	0.477	0.356
Pijnacker	0.475	0.318
Valkenburg	0.475	0.380
Martenshoek	0.474	0.374
Oss West	0.474	0.380
Purmerend	0.472	0.343
Heiloo	0.471	0.372
Etten-Leur	0.470	0.350
Ermelo	0.465	0.323
Rotterdam Noord	0.460	0.375
Arnhem Presikhaaf	0.455	0.353
Bunde	0.450	0.353
Maarn	0.444	0.334
Leiden Lammenschans	0.441	0.385
Meppel	0.433	0.330

Station name	RSQIdept	RSQIdest
Harde 't	0.432	0.320
Santpoort Noord	0.431	0.359
Vught	0.430	0.346
Sauwerd	0.425	0.330
Apeldoorn	0.417	0.341
Rheden	0.416	0.350
Rosmalen	0.415	0.325
s' Hertogenbosch Oost	0.412	0.324
Meerssen	0.409	0.328
Overveen	0.409	0.296
Driehuis	0.406	0.290
Wijchen	0.406	0.347
Heerhugowaard	0.405	0.357
Alkmaar Noord	0.401	0.407
Soest	0.401	0.270
Velp	0.399	0.320
Helmond Brouwhuis	0.397	0.320
Zaltbommel	0.396	0.295
Soestdijk	0.395	0.197
Helmond 't Hout	0.392	0.328
Soest Zuid	0.390	0.271
Nunspeet	0.390	0.290
Harderwijk	0.387	0.293
Buitenpost	0.386	0.299
Rotterdam Wilgenplas	0.386	0.266
Geleen-Lutterade	0.385	0.309
Beilen	0.385	0.316
Almelo de Riet	0.381	0.312
Echt	0.380	0.296
Wezep	0.378	0.294
Susteren	0.375	0.297
Purmerend Overwhere	0.371	0.290
Maassluis West	0.369	0.411
Rotterdam Kleiweg	0.368	0.256
Nijmegen Dukenburg	0.367	0.312
Brummen	0.365	0.295
Wierden	0.358	0.299
Duiven	0.356	0.263
Heeze	0.354	0.255
Berkel en Rodenrijs	0.353	0.273
Heino	0.350	0.257
Geleen Oost	0.350	0.284
Borne	0.348	0.279
Zevenaar	0.345	0.258
Geldrop	0.342	0.230
Oisterwijk	0.341	0.282
Cuijk	0.339	0.281
Heemskerk	0.339	0.240
Hoorn	0.338	0.319
Rotterdam Hofplein	0.331	0.220

Table 5AI. 3: Service quality indices of railway station in the Dutch national railway network (Continued)

Station name	RSQIdept	RSQIdest
Dordrecht Zuid	0.329	0.287
Lage Zwaluwe	0.327	0.275
Zuidbroek	0.323	0.254
Nijverdal	0.322	0.249
Kruiningen-Yerseke	0.316	0.313
Hoogeveen	0.313	0.258
Bergen op Zoom	0.313	0.281
Goes	0.313	0.302
Enschede Drienerlo	0.312	0.262
Goor	0.311	0.267
Hurdegaryp	0.308	0.228
Reuver	0.307	0.270
Veenwouden	0.305	0.189
Groningen	0.301	0.275
Maastricht	0.300	0.261
Dordrecht Stadspolders	0.299	0.245
Lelystad Centrum	0.298	0.242
Groningen Noord	0.296	0.242
Zandvoort aan Zee	0.294	0.220
Hoogezand-Sappemeer	0.294	0.370
Boxmeer	0.294	0.349
Veenendaal West	0.290	0.235
Steenwijk	0.290	0.200
Didam	0.288	0.244
Heerlen	0.288	0.262
Raalte	0.287	0.221
Maastricht Randwyck	0.285	0.282
Tiel	0.285	0.170
Deurne	0.283	0.306
Rotterdam Bergweg	0.282	0.191
Stedum	0.279	0.219
Kampen	0.277	0.183
Anna Paulowna	0.275	0.274
Lochem	0.272	0.197
Veenendaal Centrum	0.265	0.206
Nijmegen Heyendaal	0.265	0.224
Schagen	0.265	0.250
Bedum	0.264	0.333
Sliedrecht	0.263	0.221
Hoorn Kersenboogerd	0.262	0.217
Delden	0.259	0.220
Dronrijp	0.255	0.185
Veenendaal-de Klomp	0.254	0.190
Middelburg	0.253	0.224
Venlo	0.252	0.251
Waddinxveen Noord	0.251	0.223
Breda Prinsenbeek	0.250	0.152
Boskoop	0.250	0.224
Leeuwarden	0.248	0.227
Klarenbeek	0.248	0.186
Waddinxveen	0.238	0.209

Station name	RSQIdept	RSQIdest
Obdam	0.235	0.252
Wehl	0.234	0.220
Mantgum	0.232	0.176
Heerenveen	0.231	0.193
Leeuwarden Camminghaburen	0.228	0.177
Vlissingen Souburg	0.227	0.204
Enschede	0.222	0.209
Kropswolde	0.218	0.169
Hengelo Oost	0.218	0.187
Hardinxveld-Giessendam	0.217	0.210
Vorden	0.211	0.209
Hoogkarspel	0.210	0.186
Oldenzaal	0.209	0.166
Hoek van Holland Haven	0.208	0.183
Venray	0.207	0.185
Scheemda	0.206	0.167
Ommen	0.206	0.244
Rhenen	0.202	0.155
Doetinchem de Huet	0.200	0.169
Hoek van Holland Strand	0.197	0.168
Loppersum	0.196	0.156
Houthem-St. Gerlach	0.196	0.206
Barneveld Centrum	0.194	0.147
Wolfheze	0.194	0.155
Barneveld Noord	0.193	0.139
Terborg	0.190	0.162
Coevorden	0.185	0.177
Vlissingen	0.185	0.161
Den Helder Zuid	0.185	0.182
Olst	0.183	0.151
Hardenberg	0.182	0.180
Franeker	0.180	0.131
Landgraaf	0.176	0.144
Gorinchem	0.176	0.150
Varsseveld	0.175	0.146
Bovenkarspel-Grootebroek	0.175	0.159
Klimmen-Ransdaal	0.174	0.149
Appingedam	0.174	0.138
Zwaagwesteinde	0.173	0.140
Oosterbeek	0.170	0.134
Grijpskerk	0.170	0.134
Lunteren	0.164	0.141
Ede Centrum	0.163	0.134
Doetinchem	0.162	0.159
Ruurlo	0.162	0.159
Wijhe	0.159	0.150
Horst-Sevenum	0.154	0.131
Harlingen	0.152	0.107
Eijsden	0.149	0.127
Aalten	0.148	0.121
Schin op Geul	0.146	0.126

Table 5AI. 4: Service quality indices of railway station in the Dutch national railway network (Continued)

Station name	RSQIdeprt	RSQIdest
Zevenbergen	0.146	0.108
Oudenbosch	0.146	0.108
Sneek Noord	0.145	0.108
Blerick	0.144	0.127
Bovenkarspel Flora	0.144	0.135
Voerendaal	0.143	0.118
Sappemeer Oost	0.143	0.128
Winterswijk	0.142	0.150
Nuth	0.140	0.103
Rilland-Bath	0.138	0.134
Grou-Jirnsum	0.137	0.108
Wolvega	0.137	0.111
Lichtenvoorde-Groenlo	0.136	0.152
Sneek	0.135	0.103
Deventer Colmschate	0.134	0.111
Spaubeek	0.134	0.097
Baflo	0.134	0.105
Arnemuiden	0.131	0.119
Winschoten	0.127	0.104
Hoensbroek	0.126	0.093
Winsum	0.126	0.102
Dalfsen	0.126	0.120
Kapelle-Biezelinge	0.124	0.114
Enkhuizen	0.121	0.113
Zetten-Andelst	0.120	0.093
Akkrum	0.120	0.103
Delfzijl West	0.120	0.102
Krabbendijke	0.119	0.116
Den Helder	0.117	0.116
Schinnen	0.117	0.087
Nieuweschans	0.116	0.097
Deinum	0.114	0.079
Swalmen	0.114	0.107
Delfzijl	0.112	0.096

Station name	RSQIdept	RSQIdest
Holten	0.112	0.105
Mariënberg	0.111	0.136
Beesd	0.110	0.086
Chevremont	0.108	0.079
Eygelshoven	0.107	0.107
Tegelen	0.105	0.096
Rijssen	0.105	0.095
Harlingen Haven	0.100	0.077
Arkel	0.100	0.084
Vierlingsbeek	0.099	0.109
Hemmen-Dodewaard	0.095	0.084
Warffum	0.093	0.078
Leerdam	0.090	0.072
Usquert	0.090	0.073
Opheusden	0.089	0.073
Emmen	0.087	0.092
Kesteren	0.086	0.078
Kerkrade Centrum	0.078	0.077
Nieuw Amsterdam	0.078	0.078
Gramsbergen	0.078	0.083
Vriezenveen	0.076	0.062
Dalen	0.069	0.070
Daarlerveen	0.062	0.049
Uithuizen	0.061	0.052
IJlst	0.059	0.055
Emmen Bargeres	0.058	0.059
Uithuizermeeden	0.058	0.048
Vroomshoop	0.058	0.044
Geerdijk	0.055	0.037
Workum	0.046	0.043
Roodeschool	0.042	0.035
Hindeloopen	0.039	0.044
Koudum-Molkwerum	0.034	0.034
Stavoren	0.027	0.028

Chapter 6

6 Modelling the aggregate access mode and railway station choice

6.1 INTRODUCTION

Railway transport constitutes a sizable share of the daily travel made by Dutch travellers. The figures from the Central Bureau of Statistics (CBS) in 2002 reveal that railway transportation in the Netherlands accounts for about 8% of the overall passenger kilometres. This figure is one of the highest shares of railway transport in Europe and the world. In the US, the overall public transport share (which includes railway and bus services) is about 2% (U.S. DOT 2005). The modal split of passenger kilometres shares for the 15 Members States of the European Union are also given in Table 6.1. After Austria and France, railway transport in the Netherlands accounts for highest share of the total passenger kilometres. On the other hand, it is necessary to be aware that the railways' share in the number of trips is considerably lower, since railway trips tend to be much longer than those of other modes.

Once the decision to travel by train is made, some of the logical questions that follow are: 1) Which station to use for departure?; 2) Which access mode to use to get to the station?; and 3) Which route to follow to the destination? The decisions on these types of choice are affected by different factors. Bovy and Stern distinguish three factors: 1) features of the available alternatives; 2) characteristics of the traveller; and 3) features of the choice situation (Bovy and Stern 1990). This chapter is a study on the first two types of choice facing railway travellers mentioned above. These are the choice of departure railway station, and the choice of access mode to the railway station. The chapter does not attempt to address the issue of route choice to a destination. The trip from the origin to the departure station is called the *access* part, while the trip from the destination station to the final destination is called the *egress* part. This chapter addresses the two basic choices made by railway travellers concerning the access part of a train trip: access mode and departure station choices. Therefore, the features that are included in our analysis are selected from the access point of view. Decisions on these choices are expected to be based on the assessment of a number of relevant features.

The choice of a departure station is influenced by two types of features: features related to the accessibility of the station, and features related to the rail services provided at the station. Easily accessible railway stations are more likely to be selected as a departure station than less accessible stations. For instance, keeping other things constant, stations served by frequent public transport modes are expected to be preferable to stations which have less-frequent public transport services as departure stations. Similarly, the availability of other access modes such as car, public transport, and other non-motorized modes is expected to influence the choice of a departure railway station. Moreover, the choice of a departure station also depends on the quality of the station itself. The quality of a railway station is generally explained by the quality of rail and supplementary services provided at the station. The frequency of train services, network connectivity, and coverage are some examples of the rail service. The presence of other supplementary facilities such as the availability of parking spaces, the park-and-ride possibility, bike stands and storage facilities (lock-ups) also boost the attractiveness of a station as a departure station. In the previous chapter, we discussed the pure rail service quality (RSQI) measure of a railway station.

Table 6.1: Modal split by country for passenger transport (in passenger kilometres share): EU-15 (5 modes) in 2002

	CAR	BUS	RAILWAY	TRAM & METRO	AIR
BELGIUM	79.8	9.9	6.0	0.7	3.6
DENMARK	74.3	11.1	6.8		7.8
GERMANY	78.8	8.6	7.8	0.9	3.9
GREECE	65.9	17.0	1.4	1.0	14.6
SPAIN	71.2	10.6	4.5	1.2	12.5
FRANCE	83.1	4.5	8.2	1.2	3.0
IRELAND	72.8	12.4	3.2		11.5
ITALY	80.2	11.0	5.3	0.6	3.0
LUXEMBOURG	74.7	12.8	5.1		7.4
NETHERLANDS	81.5	4.1	8.1	0.8	5.5
AUSTRIA	70.7	13.6	8.4	2.8	4.5
PORTUGAL	79.7	8.3	3.1	0.5	8.3
FINLAND	77.7	10.3	4.4	0.7	7.0
SWEDEN	74.0	8.0	7.2	1.8	9.0
UNITED KINGDOM	80.9	5.9	5.1	1.1	7.1

Source: Adapted from *EU energy and transport in figures: statistical pocket book 2004*

The revealed choice data for departure stations for Dutch railway travellers shows that, in about 47% of the cases, passengers choose a departure station which is not the nearest station to their places of residence. This indicates that the measurement of distance to the railway station for measuring railway accessibility has some limitations. This method indirectly assumes that railway stations are identical to each other, except for the distance from the location of the user to the railway station. However, in reality, railway stations differ from each other in many respects. In the literature, the typical distinction between railway stations is made over the type of the station. Four types of railway stations can be identified: commuter railway stations; heavy railway stations; light railway stations; and bus rapid transit (BRT) stations (*see* Chapter 2). Even with such distinctions we still observe heterogeneity among stations of the same type. Thus, there is a need for a comprehensive method for distinguishing the features of railway stations for a proper analysis regarding railway accessibility and departure station choice. One way of arriving at this measure is to understand the decision process for using a departure station. Thus, the first step is to distinguish the features that have railway accessibility implications. In this context, these factors can be summarized as follows. First, the ease of reaching the station plays an important role in determining the accessibility of a station. Distance from the origin to the departure station can be taken as a general proxy. In addition, because accessing the railway station can be done by different modes of transport, mode-related features are also important factors in the determination of the ease of accessing the railway station. Features related to the quality of road access and public transport can be mentioned. Supplementary station services such as the availability of parking space and bike stands also contribute to the access mode choice. The second component relates to the level of rail service that is delivered at the railway station. This was the subject matter of Chapter 5 of this thesis.

By applying these railway station accessibility concepts, this chapter aims to analyse the choice process of Dutch travellers for access mode and departure stations. This will, in turn, be used to calculate a general railway accessibility index for zones where people live. In most real estate price studies, railway station accessibility is just given by the distance to the nearest railway station from the property in question. However, railway station accessibility encompasses all aspects that are involved in the choice process for a departure station. The accessibility of a railway station can thus be considered to encompass all the features that travellers consider in their choice of a departure station. This method of calculating an accessibility index is also expected to single out the pure railway transport-related effect. Thus, this index is considered to be superior to previous methods. Furthermore, understanding the valuation and decision mechanism leading to the choices of departure station and access mode has several practical implications for the formulation of transportation management policy for urban areas. In the first place, it enables us to define the catchment areas (market areas) of the stations. This means that it enhances the predictions of travel demand at station level. This in turn can be used as a basis for site selection for the development of new lines or planning extensions for existing lines, as well as parking facilities and feeder public transport operation planning. In addition, the understanding of the sensitivity of travellers towards the access and station features gives a station operator the basis for increasing travellers' turnover.

The chapter is organized as follows: Section 6.2 briefly reviews the literature in the area. Section 6.3 gives the specification of the nested logit model which is applied in the estimations of this chapter. In Section 6.4, we discuss the specification of the utility models for the access-mode departure station choice. In Section 6.5, we describe the data used in our analysis. Section 6.6 gives the estimation results, followed by the discussion of these results. Section 6.7 ends the chapter with summaries and conclusions.

6.2 LITERATURE REVIEW

The literature on access mode and departure station choice is generally limited. One of the early rail transit station choice models was developed by Kastrenakes (1988) in an effort to prepare a basis for forecasting railway travel in the New Jersey area. With origin-destination pair data, he analysed the choice process for a departure station by considering of the access time required to reach the station, the frequency of service at the boarding station, whether the boarding station is located in the locality of the passenger's residence, and the generalized cost of the train trip between the departure station and the destination station (Kastrenakes 1988). The study found, as expected, positive effects for frequency of service and location of the station in the locality of the passenger's residential area on the probability of departure station choice. Similarly, the expected negative effects were found for access time and the generalized cost of the rail trip. In another study, Wardman and Whelan (1999) studied railway station choice for the London area. This study was done in relation to parking attractiveness for station choice. It was indicated that the availability of a parking area in a station and other station facilities are important features for station choice (Wardman and Whelan 1999).

Some studies on this theme have also incorporated access mode choice in a nested structure (Fan et al. 1993; Wardman and Whelan 1999; and Davidson and Yang 1999). Generally, the access mode choice at the upper-level of the nest was the accepted structure rather than the reverse order. Fan et al. (1993) included several variables for the transit station choice. Travel time (including access and in-vehicle time), fare, peak-hour frequency of trains, and the number of parking places were among the included variables. As expected it was found that the coefficients for frequency of service and parking had a positive sign and coefficients for travel time and fare had a negative sign. Wardman and Whelan (1999) on the other hand, compared the access mode-station choice for business and leisure travellers. They found the value of time is highest for business trips and lower for leisure trips. Other variables included were journey time, journey headway, facilities at the station, and parking availability. They all show expected the signs and significant effects on the choice of the departure station.

Choice analysis of this form has been popular in the literature on airport and airline choice (Ashford and Bencheman 1987; Hess and Polak 2004; Pels et al. 2001; Pels et al. 2003; and Basar and Bhat 2004). Fares (airport tax), access time, frequency of service, and other facilities are important features used in airport choice. Some studies also include time-series historic data in the choice features of those commuters who tend to keep on using an airport that they have previously used. The analyses of departure airports have some relevance to the railway station choice. Most of the time, the fare difference between railway stations are not observed. Thus, the fare does not play a relevant role in the choice among stations. However, access features like access time and access cost are obviously relevant for the railway station analysis. The frequency of service, as indicated by the number of trains leaving the station per given time interval and/or the number of destinations served directly from the station, plays an important role in station choice analysis. The same holds for the nature of the station and facilities at the station. Obviously, international and intercity stations are expected to enjoy higher choice probabilities compared with express or stop train stations[9]. Stations with better public and passenger-related facilities are also expected to be more attractive compared with stations with less or no facilities. The attractiveness of the station as a departure station declines as the access time increases.

[9] In the Netherlands there are four types of railway services: namely, the all-station rail services called ‘stop train’; ‘semi-fast’ also called ‘express’ rail services which call at main and medium cities; ‘intercity’rail services that only call at main cities; and international trains that only stop at a very limited number of stations.

In this chapter, we analyse the choice process by which Dutch households select a particular station as the origin of their trip. We examine the effect of distance, service level, and various station facilities in the underlying utility level of the choice model.

6.3 THE NESTED LOGIT MODEL

The choice of a railway station and an access mode can be assumed to be the result of a utility maximization process. Given the situation under which the choice is made, alternatives bring certain utility levels to the travellers. Passengers choose a combination of access mode and departure railway that provides a maximum implicit utility among all alternative combinations. Choice based on the relative attractiveness of competing alternatives from a set of mutually exclusive alternatives is called a discrete choice situation. The Multinomial Logit Model (MNL) is among the first models designed to model a discrete choice situation that involves several alternatives. The MNL model assumes that the unobserved utility components of alternatives are independently and identically distributed. This leads to the proportional substitution property between alternatives. That is to say, the ratio of the station choice probabilities of two alternatives is not affected by the presence or absence of other alternatives. This is generally known as the Independence from Irrelevant Alternatives (IIA) property. However, this assumption causes serious limitation to some applications. Several extensions to this model have been developed in which the IIA assumption is not necessary. The nested logit model is an extension of the multinomial logit model which is widely used to model hierarchical choice situations. It has been introduced by, amongst others, Ben-Akiva (1973), and allows alternatives to be correlated so that the IIA assumption does not hold. In the nested logit model, correlated, alternatives are assigned to the same nest. Alternatives in different nests are uncorrelated and thus appear to take place at different levels. Choices at the lower-level are called ‘elemental choices’, whereas choices at the upper-levels are called ‘structural choices’. The ordering of the choices in the decision tree pertains to the grouping of similar (correlated) choices rather than to the sequence of the decisions. The underlying assumption of the process is, however, that decisions are taken simultaneously rather than sequentially. Thus, the nested logit model presents a generalized situation of the multinomial logit model by allowing dependence between the error terms of similar choices. Dependency between choices can occur at different levels. The level of dependency thus determines the level at which the choice should be placed in the nest. In this chapter we apply the nested logit model to model the choices made concerning departure railway station and access mode to the departure railway station by Dutch railway travellers. Thus, our choice analysis has two levels. There are two possible decision structures, depending on which choice determines the nest. We will analyse both structures to determine which nest is appropriate to model the choice behaviour. Next we will specify the econometric model.

6.3.1 The econometric model

Let us assume that the decision structure has two levels. There are K alternatives which can be grouped into J nests, in which each nest has N_j alternatives. The final choice can be regarded as a choice concerning the combination of choices on both levels. Suppose the utility of the final choice for the choice maker is:

$$U_{jk} = V_{jk} + \varepsilon_{jk}, \tag{1}$$

where, V_{jk} is the systematic utility of the final choice; and ε_{jk} is the non-systematic part of the utility for the final choice. If we assume that ε_{jk} are *iid* Gumbel extreme-value distributed, the probability of the outcome can be given by the logit model:

$$P_{jk} = \frac{\exp(V_{jk})}{\sum_{m \in J} \sum_{l \in K_m} \exp(V_{ml})}. \tag{2}$$

Let us further assume that the utility is a linear function of the features of the choice nodes. The utility function of an alternative is composed of two parts: a part specific to the alternative, and a part associated with the nest. Thus, the total systematic utility of the final choice can be given by;

$$V_{jk} = V_{k|j} + V_j = \boldsymbol{\beta}' \mathrm{x}_{k|j} + \boldsymbol{\gamma}' \mathrm{y}_j. \tag{3}$$

where, x and y are features related to the elemental and structural choices, respectively; and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are the corresponding coefficients. However, the choice of a nest is expected to be based on the expected utility which includes the inclusive value from alternatives within the nest. We impose the scaling parameters (μ_j) at the elemental level, and normalize the scaling parameter at the nest level to unity. The expected systematic utility $(\tilde{V})$ at the nest level is given by:

$$\tilde{V}_j = V_j + \frac{1}{\mu_j} \ln \sum_{l \in K_j} \exp(\mu_j V_{l|j}) . \tag{4}$$

This normalization option of the scaling parameters is generally referred as RU2 (Hensher and Greene 2002). The logit models based on RU2 are consistent with Random Utility Maximization theory when the scaling parameters (μ_j) are greater than 1. Thus, the probabilities for both the elemental and the structural choices can be given as:

$$P_{k|j} = \frac{\exp(\mu_j V_{k|j})}{\sum_{l \in K_j} \mu_j V_{l|j}} = \frac{\exp(\mu_j \boldsymbol{\beta}' \mathrm{x}_{k|j})}{\sum_{l \in K_j} \exp(\mu_j \boldsymbol{\beta}' \mathrm{x}_{l|j})} = \frac{\exp(\mu_j \boldsymbol{\beta}' \mathrm{x}_{k|j})}{\exp(I_j)} , \tag{5}$$

where,

$$I_j = \log(\sum_{k \in K_j} \exp(\mu_j \boldsymbol{\beta}' \mathrm{x}_{k|j}) , \tag{6}$$

$$P_j = \frac{\exp(\tilde{V}_j)}{\sum_{m \in J} \exp(\tilde{V}_m)} \frac{\exp(\boldsymbol{\gamma}' \mathrm{y}_j + 1/\mu_j (I_j))}{\sum_{m \in J} \exp(\boldsymbol{\gamma}' \mathrm{y}_m + 1/\mu_m (I_m))} . \tag{7}$$

Applying probability theorem $(P_{jk} = P_{k|j} \times P_j)$, the joint probability is given by:

$$P_{jk} = \frac{\exp(\mu_j \boldsymbol{\beta}' \mathrm{x}_{k|j})}{\exp(I_j)} \times \frac{\exp(\boldsymbol{\gamma}' \mathrm{y}_j + 1/\mu_j (I_j))}{\sum_{m \in J} \exp(\boldsymbol{\gamma}' \mathrm{y}_m + 1/\mu_m (I_m))} . \tag{8}$$

This equation can be estimated using maximum likelihood. $1/\mu_j$ is known as the 'inclusive value parameter' since it is the estimated coefficient of I_j. It can be interpreted as a measure of dissimilarity between alternatives within a nest. It is an indicator of the correlation in the unobserved components of the utilities of the choices grouped under nest j. The smaller the value of the inclusive value parameter, the higher is the correlation between the alternatives in the nest. It can also be shown that $(1 - 1/\mu_j^2)$ is equal to the correlation of the utilities of alternatives within nest j (Ben-Akiva and Lerman 1985). If $\mu_j = 1$, the situation is characterized by complete independence among the alternatives in the nest. This suggests that there is no need for grouping the alternatives in nests, and thus the nested logit model collapses into the multinomial logit model.

6.3.2 Overall utility

The overall utility level that a traveller assumes is determined by the utility level she or he enjoys by making choices on access mode and departure railway station over her or his choice set. The overall utility level is equal to the inclusive value of utility at the choice maker level. On the basis of the above model specification, the inclusive utility level determining the overall railway accessibility level is given as follows:

$$IV_p = \log \sum_{j=1}^{J} \exp\left[\boldsymbol{\gamma}' \mathrm{y}_j + \frac{1}{\mu_j} I_j \right].$$

The unit of analysis in this study is a postcode area. Therefore the inclusive value represents the overall utility level that a postcode area enjoys in relation to railway travel. This measure is expected to provide a richer measure of railway accessibility compared with previous ways of measuring railway accessibility. In addition, it provides a flexible way of comparing the effect of change in the underlying components. For example, we can easily compare the effects of changes in public transport settings and railway services at stations on the overall railway accessibility. Most of the time, the decisions on these aspects are the responsibilities of different parties. Thus, the measure of overall accessibility gives an opportunity to integrate decisions of different parties toward a shared goal. In Chapter 7 we will use indices based on this measure to represent general railway accessibility in the house price estimations.

6.4 UTILITY SPECIFICATION

We start with the assumption that the passenger in our analysis has already decided to travel by train. The passenger then faces two related choices: 1) the choice of the access mode ($a \in A$) to take in order to reach a station; and 2) the choice of the departure station ($d \in D$). Both choices are made simultaneously: travellers choose a combination of access mode and a departure railway station. We distinguish three possible choice structures. In the first structure, we consider groups (nests) of alternatives with a common access mode; the alternatives within the nest are correlated. In the second structure, we consider groups (nests) with the departure station as a common element. In the third structure, there is no common element: all alternatives are independent. The last structure is basically modelled by the multinomial logit model. However, the first two structures are modelled by the nested logit model. The appropriateness of a certain choice structure depends on the assessment of which one results in a sound grouping. As mentioned earlier, the grouping of alternatives within nests is motivated by the dependency in the error (unobserved) components of the utility of similar choices. Apart from intuitive subjective judgments, there is no indicator to say which tree structure is appropriate at the start. However, the inclusive value parameters given by the estimation give an indication of which nest structure is appropriate for modelling the choice analysis. We will discuss this in detail later in this section. Next, we will discuss the specifications of the utility model for both cases of the nested structure.

6.4.1 Access mode- departure station

This nest structure puts access mode in the upper-level and the choice of a departure station at the lower-level. This structure is motivated by the fact that the unobserved components of station utilities accessed by the same mode of transportation are correlated. The decision tree for this choice can be depicted by Figure 6.1 below.

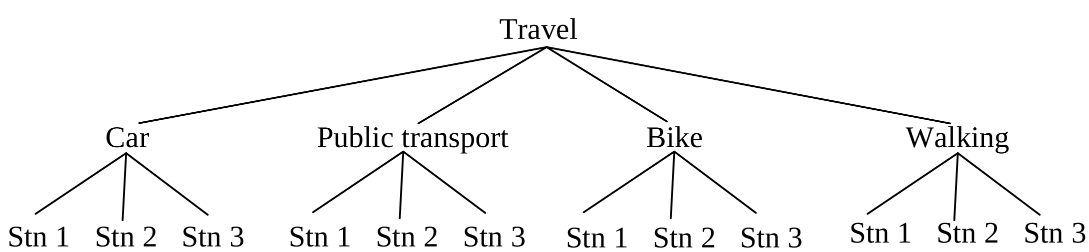


Figure 6.1: Access mode departure station choice decision tree

Let us start by specifying the utility of the branch (access mode) level. We assume a linear functional form for the underlying utilities. The underlying utility function of the access mode choice is a function of features at the postcode area level and the inherent characteristics of the access mode. From our data, the car ownership (number of cars per person) level in the postcode area is the only feature that is related to the mode choice and which is not linked to the departure station choice. The systematic utilities of the four access modes are given below (see Equations 9-12).

$$V(car) = \alpha_{car} + \beta_{car\_carown} * carownership\,; \tag{9}$$

$$V(PT) = \alpha_{PT} + \beta_{PT\_carown} * carownership\,; \tag{10}$$

$$V(bike) = \alpha_{Bk} + \beta_{Bk\_carown} * carownership\,; \tag{11}$$

$$V(walk) = 0\,. \tag{12}$$

where, $V(.)$ gives the systematic utility of the access modes; $carownership$ is the level of car ownership in the postcode area; the βs are the coefficients for the effect of car ownership on the utility level of corresponding access modes. The mode-specific constants account for the mode related characteristics. The effect of the car ownership effect on the choice of access modes is expected to explain the substitution/competition effect between car as an access mode and the other alternative modes. A positive coefficient for car ownership implies that an increase in car ownership promotes the use of the specified mode. On the other hand, a negative coefficient implies that an increase in car ownership discourages the use of the specified mode. We expect an increase in the car ownership level in the postcode area to promote the use of the car access mode and to discourage the use of the other access modes. However, the negative effect is expected to be more intense on the longer-distance-oriented motorized mode: namely, public transport, than on bike and walking. The walking mode is set to serve as a reference point for the other modes.

The lower-level choice relates to the departure station choice. The station choice utilities are assumed to be determined by characteristics related to the stations and characteristics linking the access mode and the stations. Thus, we adopt a generic utility formulation for the departure-station choice quality. Differentiations are only made on the basis of which mode is used to access the stations. The station characteristics are given by the rail service quality index (RSQI) determined in Chapter 4. Even though the distance variable enters the systematic utilities of the stations, it may have a different implication for the departure-station utility based on the access modes applied. Thus, we differentiate the effect of distance to the station by access mode. Distance is expected to have a negative effect on the utility in all four cases. However, the magnitude of the effect is expected to be higher for short-distance-oriented modes than long-distance-oriented modes. A higher negative effect of distance is expected for walking and bike modes than for public transport and car modes. The utilities attached to each of the access modes are expected to decline with distance. In general, the RSQI of the station is expected to have a positive impact on the utilities of the departure station accessed by all modes. The presence of supplementary station facilities are also expected to be access-mode-related. For instance, the presence of a parking area at the station is only expected to affect the utility of departure stations accessed by the car mode. It is expected to have a positive effect on the utility of stations accessed by car mode. Similarly, a bicycle stand is expected to influence the choice of station accessed by bike. A positive effect is expected. Public transport travel time and frequency influence the choice of departure railway station accessed by public transport. The frequency of public transport is expected to have a positive effect on the utility of departure stations accessed by public transport. On the other hand, public transport travel time is expected to negatively affect the utility of the stations accessed by public transport. The systematic utility functions of a departure station choice, given an access mode, are specified as follows:

$$V(station_k \mid car) = \beta_{cardist} \times dist_k + \beta_{RSQI} \times RSQI_k + \beta_{parkcar} \times parking_k; \tag{13}$$

$$\begin{aligned} V(station_k \mid pt) = \beta_{PTdist} \times dist_k + \beta_{RSQI} \times RSQI_k + \beta_{PTtravltime} \times PTtraveltime_k \\ + \beta_{PTfreq} \times PTfreq_k; \end{aligned} \tag{14}$$

$$V(station_k \mid pt) = \beta_{bikedist} \times dist_k + \beta_{RSQI} \times RSQI_k + \beta_{\mathrm{bikestand}} \times \mathrm{bikestand}_k; \tag{15}$$

$$V(station_k \mid walk) = \beta_{walkdist} \times dist_k + \beta_{RSQI} \times RSQI_k. \tag{16}$$

where, $k \in K = \{1,2,3\}$ is an element of the set of departure stations for the postcode area; $dist$ is the distance from the centroid of the postcode area to the railway station considered; $RSQI$ is the rail service quality index; $parking$ is a dummy variable indicating the presence of a parking area in or around the railway station; $PTtraveltime$ is the average public transport travel time from the postcode area to the railway station given in minutes; $PTfreq$ is the average public transport frequency of service from the postcode area to the railway station given by the number of services per hour; and *bikestand* is a dummy variable indicator for the presence of bicycle stand at the railway station.

6.4.2 Departure station-access mode

An alternative way of arranging the choices concerning departure station and access mode is to put the departure station on the upper-level and the access mode choice at the lower-level of the nest, as depicted by Figure 6.2. This grouping assumes that there are similarities between access modes that are used to access the same departure station.

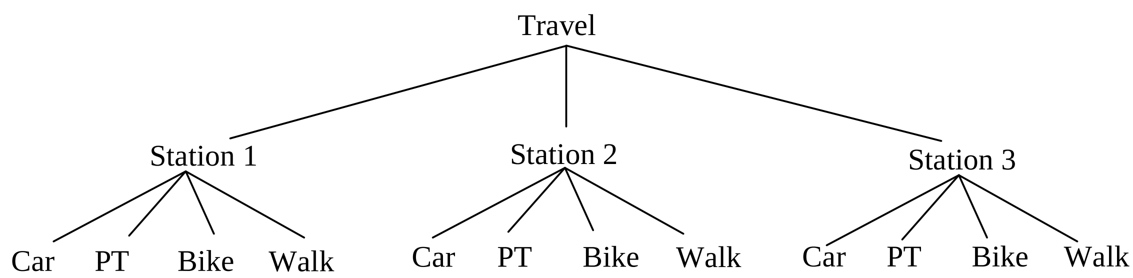


Figure 6.2: Departure station–access mode choice decision tree.

We assume the utilities of the upper-level choice alternatives: namely, concerning the choice of a departure station, are affected by the RSQI determined earlier in Section 5.2. The generic departure station utility function is given Equation 17 below:

$$V(station_k) = \beta_{RSQI} \times RSQI_k \tag{17}$$

The utility functions for the access mode are explained by a number of variables. To account for the mode-specific effects, the functions include the corresponding mode-specific coefficients. No prior expectations are made on the sign or magnitude of the coefficients. Car ownership levels are expected to affect the utility of all modes. The inclusion of car ownership in the utility specification is aimed at capturing the competition effect. As previously discussed, the walking mode is taken as the reference group. The utility specifications for the access modes also include the distance and station features that are related to the specific access mode. The distance effect is assumed to be mode-specific. Some railway station features are also expected to affect the utilities of certain access modes, and not others. For instance, the availability of a parking area in or around the station is related to car access mode. Similarly, the presence of a bicycle stand at the station is a feature related to the bicycle access mode. The specifications for the different access-mode choice utilities, given that station k is chosen as a departure station, are given by Equations 18-21. The variables are explained in the previous subsection.

$$V(car \mid station_k) = \alpha_{car} + \beta_{car\_carown} \times carownership + \beta_{cardist} \times dist_k + \beta_{parkcar} \times parking_k; \tag{18}$$

$$V(pt \mid station_k) = \alpha_{PT} + \beta_{PT\_carown} \times carownership + \beta_{PTdist} \times dist_k + \beta_{PTtravltime} \times PTtraveltime_k + \beta_{PTfreq} \times PTfreq_k; \tag{19}$$

$$V(bike \mid station_k) = \alpha_{BK} + \beta_{BK\_carown} \times carownership + \beta_{bikedist} \times dist_k + \beta_{\mathrm{bikestand}} \times \mathrm{bikestand}_k; \tag{20}$$

$$V(walk \mid station_k) = \beta_{walkdist} \times dist_k. \tag{21}$$

6.5 DATA

The data used in our analysis were obtained from the Dutch National Railway Company (Nederlandse Spoorwegen–NS). A postcode area is the unit of analysis. Household choices for access mode and departure railway station are aggregated at this level of zoning. The final analysis is based on 1440 postcode areas. For each of the postcode areas, a set of three departure stations is identified. In most cases the set accounts for the three most frequently-used departure railway stations in the postcode area. In some other cases, the set is determined on the condition of proximity to the centroid of the postcode area. The set of departure stations for each postcode area are ranked according to the size of the share of usage they account for as departure stations. This means the first station accounts for the highest share of usage as a departure station in the postcode area, whereas the third station accounts for the least of the three. The sum of the shares accounted for by the three stations in each postcode area constitutes 100% of the departure station usage. In total 346 railway stations are included in the analysis. In addition, a set of four alternative modes is defined for each postcode area: car, public transport, bike and walking. All four access modes are assumed to be available for each postcode area. All choices are given in shares because of the aggregated nature of the data. Thus, the final choice explains the joint share of access-mode and departure station choices made in the postcode areas. Each postcode area faces 12 access mode and departure-station choice combinations.

The data set incorporates several features related to the railway stations and access modes. The car ownership level is one of the relevant features given at the postcode area level. At the station level, we find data for the RSQI, availability of parking areas, and availability of bike stands. Public transport data on frequency and travel time were retrieved from the public transport timetables of the lines linking the postcode area and each of the alternative departure stations. The public transport timetables are available at the 6-digit postcode level – an area comprising up to about 50 houses, and were aggregated to the 4-digit postcode level – an area composed of about seven 6-digit areas. GIS information on the location of the centroid of the postcode area and the railway stations was used to determine the distance measure to represent the accessibility indicator. Thus, our data set includes the usage share of the three most frequently-chosen stations for each postcode area and the railway station features of each station including the distance between the centroid of the postcode area and the railway station.

Description of Station and Access-mode Characteristics

As has been previously discussed in this chapter, we assume railway station accessibility is explained by two factors: the ease of reaching the stations, and the service levels provided at the stations. The ease of reaching the stations is linked to the distance between the departure point (the centroid of the postcode area in this case) and the railway station. On the other hand, the level of services provided at the stations is related to the frequency of trains leaving the station per period of time and network connectivity, as determined by the number of destinations that can be reached without transfer. The RSQI (*see* Chapter 5) is determined from the generalized journey time between stations; the importance of the destination station; and the ratio of the generalized journey time to the distance. The attractiveness of a station can also be affected by facilities that supplement railway transport. Parking areas, the availability of a park-and-ride facility, and bike stands can be mentioned. The choice probabilities of access-mode and departure-station in the postcode areas are summarized in Table 6.2 below. It is based on the access mode – departure station tree structure of choices. Public transport, with about 38% of the share, is the most frequently used access mode by which passengers reach the railway station. On the departure station side, on average the first most frequently-chosen railway station accounts for about 77% of the total share. The second and third most frequently-chosen railway stations account for 17% and 6%, respectively.

Table 6.2: Summaries of choice probabilities

Access mode	Departure station	Branch level	Choice level
Car	1st Station	0.2376	0.1596
	2nd Station		0.0527
	3rd Station		0.0253
Public transport	1st Station	0.3765	0.2862
	2nd Station		0.0680
	3rd Station		0.0223
Bicycle	1st Station	0.2443	0.2056
	2nd Station		0.0309
	3rd Station		0.0078
Walking	1st Station	0.1416	0.1220
	2nd Station		0.0162
	3rd Station		0.0034

Table 6.3 below gives the descriptive statistics of railway station characteristics and the accessibility indicators for the postcode areas. For the purpose of showing the variation in Table 6.3, we only give the statistic on the distance of the most frequently-chosen station from the postcode area. In addition, Table 6.3 gives the railway station features. Included are the indicators of RSQI and supplementary facilities; frequency and travel times of public transport service; car ownership level in the postcode areas; and distance measure to railway stations.

Table 6.3: Descriptive statistics for the railway station characteristics (2001/2002)

Description	Number of stations/ postcode areas	Min	Max	Mean	Std. Deviation
Rail service quality index (RSQI)	365	0.03	2.00	0.43	0.33
Bicycle stand	96			0.28	
Parking	318			0.91	
Accessibility from postcode areas					
Distance to the most frequently-chosen station (m)	1400	95	31,708	5,840	5,583
Car ownership in the postcode area	1440	0.11	0.99	0.40	0.09
Frequency of public transport (vehicle per hour)	1440	1.00	19.00	1.98	2.14
Public transport travel time (minutes)	1440	2.00	57.91	25.41	12.81

6.6 ESTIMATION AND DISCUSSION

The estimation results of the nested logit model for the two nest structures discussed above are given in Tables 6.4 and 6.5 below. The inclusive value parameters in the estimations give us an indication as to which nesting structure is more appropriate for modelling the choice behaviour. The two estimations are readily comparable since they both use the same normalization procedure for the scaling parameters in the model. The scaling parameter is normalized at the upper-level, and the lower-level scaling parameters are free. This model is generally referred to as the 'Random Utility Model 2' (RU2). For the model outcome to be consistent with random utility maximization, the inclusive value parameters should be greater than 1. A value which is equal to 1 indicates a complete collapse of the nested logit model as a multinomial logit model. Generally speaking, most variables in the estimations have significant and expected effects. However, the inclusive value parameters in the departure-station – access-mode nest structure fall below 1. This indicates that this structure is not appropriate for nesting the choices. On the other hand, the inclusive value parameter estimates based on the access-mode – departure-station choice structure are above 1. Thus, this nesting structure seems more appropriate for the choices than the reverse order nest. Our discussion will, therefore, focus on the estimation result of the access-mode – departure-station nest structure.

Table 6.4: Estimation results for access-mode – departure-railway station decision nest (RU2)

Variable	Coefficient	Z-value	P-value
Lower-level parameters			
RSQI	1.0654	8.652	0.000
CAR_DIST	-0.1088	-8.539	0.000
PARK_CAR	0.9348	2.777	0.006
PT_DIST	-0.0472	-4.506	0.000
PT_FREQ	0.1057	5.148	0.000
PT_TIME	-0.0108	-2.275	0.023
BK_DIST	-0.4833	-13.643	0.000
BIKE_STAND	0.3800	3.737	0.000
WK_DIST	-1.1222	-13.030	0.000
Upper-level parameters			
ALPHA_CAR	-3.7989	-6.608	0.000
CAR_CAROWN	0.7536	0.702	0.483
ALPHA_PT	-0.8643	-2.035	0.042
PT_ CAROWN	-4.2328	-4.512	0.000
ALPHA_BIKE	-1.0871	-2.735	0.006
BK_ CAROWN	0.3372	0.359	0.720
Inclusive value parameters (μ)			
CAR	1.628	10.995	0.000
PUBLICT	1.628	10.995	0.000
BIKE	1.628	10.995	0.000
WALKING	1.628	10.995	0.000
number of observations	= 17280		
log likelihood function	= -2678.118		
Restricted log likelihood	= -3578.266		
Chi squared	= 1800.295		
Degrees of freedom	= 16		
Prob[ChiSqd > value]	= 0.0000		
R-sqrd	= 0.25156		
RsqAdj	= 0.25080		

Table 6.5: Estimation results for departure railway station– access -mode decision nest (RU2)

Variable	Coefficient	Z-value	P-value
Lower-level parameters			
ALPHA_CAR	-7.754	-5.820	0.000
CAR_CAROWN	1.545	0.776	0.438
CAR_DIST	-0.150	-8.161	0.000
PARK_CAR	2.040	2.259	0.024
ALPHA_PT	-1.653	-2.162	0.031
PT_CAROWN	-8.523	-4.595	0.000
PT_DIST	-0.029	-1.596	0.111
PT_TIME	-0.022	-2.202	0.028
PTFREQ	0.225	5.608	0.000
ALPHA_BIKE	-2.347	-3.002	0.003
BK_CAROWN	0.574	0.320	0.749
BIKE_DIST	-0.878	-13.616	0.000
BIKE_STAND	1.115	4.798	0.000
WALK_DIST	-2.219	-11.196	0.000
Upper-level parameters			
RSQI	1.576	11.614	0.000
Inclusive value parameters (μ)			
STATION 1	0.495	10.364	0.000
STATION 2	0.495	10.364	0.000
STATION 3	0.495	10.364	0.000
number of observations	= 17280		
log likelihood function	= -2680.225		
Restricted log likelihood	= -3578.266		
Chi squared	= 1796.082		
Degrees of freedom	= 16		
Prob [ChiSqd > value]	= 0.00000		
R-sqrd	= 0.25072		
RsqAdj	= 0.24996		

6.6.1 Effect of station's rail service quality

The estimation results show that the measure of the rail service quality index (RSQI) has a positive and significant effect on the choice of departure stations. In addition, the presence of supplementary facilities at the stations also has a positive impact on the choice of a departure station. The presence of a parking area and bike stands have a positive and significant effect on the choice of departure stations accessed by car and bike, respectively. The elasticities of the RSQI on the choice probability of access mode and departure station are presented in Table 6.6 below.

Table 6.6: Direct elasticity of rail service quality index (RSQI)

Access mode	Departure station	Elasticity		
		Branch level	Choice level	Total elasticity
Car	1st Station	0.313	0.519	0.832
	2nd Station	0.231	0.689	0.92
	3rd Station	0.171	0.723	0.894
Public transport	1st Station	0.286	0.537	0.823
	2nd Station	0.187	0.763	0.95
	3rd Station	0.134	0.806	0.94
Bicycle	1st Station	0.37	0.28	0.649
	2nd Station	0.24	0.612	0.851
	3rd Station	0.171	0.659	0.83
Walking	1st Station	0.378	0.136	0.514
	2nd Station	0.264	0.392	0.656
	3rd Station	0.221	0.484	0.705

The choice level represents the departure stations accessed by a given access mode. The three stations accessed by a given mode are arranged according to the size of their market share in the postcode area. The 1st station is the most frequently-chosen station in the postcode area. The table shows that the elasticity of the RSQI on the choice level increases as we go from the biggest station to the smallest station accessed by all modes. For example, a 1% increase in the RSQI of railway stations accessed by the car mode leads to an increase in the choice probability of the station by 0.52%, 0.69%, and 0.72% for the 1st, 2nd, and 3rd stations, respectively. The trend of the effect is consistent with the intuitive expectation, in that an increase in a station's RSQI is expected to a have higher impact on the stations with the lowest share. Note that the elasticities given in the table represent direct elasticities. Cross-elasticities are not reported. An increase in the service quality of a railway station leads to an increase in demand. The travel demand increase experienced in one station comes at the expense of the demand loss at the other railway stations accessed by the same mode of transport and railway stations accessed by other modes. Because of the higher similarity between railway stations accessed by the same access mode, changes in the RSQI of a station are expected to result in a higher substitution between the stations within the nest than outside the nest. Thus, the cross-elasticity of rail service quality of a station is expected to be higher for a station within one nest than for stations across nests. To illustrate the effect of change in the rail service quality of a station on the choice share, let us take the case of the station with the highest share accessed by car mode. The direct and cross-elasticities of rail service quality change of the 1st station accessed by car mode are given below, in Table 6.7.

Table 6.7: Direct and cross-elasticities of rail service quality index for the station with the highest share accessed by car

Access mode	Departure station	Elasticity		
		Branch level	Choice level	Total elasticity
Car	1st Station	0.313	0.519	0.832
	2nd Station	0.201	-0.462	-0.261
	3rd Station	0.193	-0.451	-0.258
Public transport	1st Station	-0.110	0.000	-0.110
	2nd Station	-0.080	0.000	-0.080
	3rd Station	-0.082	0.000	-0.082
Bicycle	1st Station	-0.071	0.000	-0.071
	2nd Station	-0.076	0.000	-0.076
	3rd Station	-0.071	0.000	-0.071
Walking	1st Station	-0.040	0.000	-0.040
	2nd Station	-0.061	0.000	-0.061
	3rd Station	-0.057	0.000	-0.057

Based on these elasticities, we can compute the share of each station for any change in the 1st station's share accessed by car mode. For comparison purposes, we give the change of shares as a result of a 10%, 50%, and 100% increase in the rail service quality of the 1st station accessed by car. The resulting shares are given in Table 6.8 below. As the RSQI increases. the share of the 1st station accessed by car increases. This leads to a decrease in the shares of other stations. In relative terms the other stations accessed by the same mode of transport (car in this case) lose more shares than the other stations accessed by other modes of transport. This shows the close similarity of stations accessed by the same mode, which in turn facilitates substitution between stations in the event of changes in the underlying features.

Table 6.8: The effect of change in the RSQI of the station with the biggest share accessed by car

Access mode	Departure station	Base shares	Share after 10% increase in the RSQI	Share after 50% increase in the RSQI	Share after 100% increase in the RSQI
Car	1st Station	0.160	0.173	0.227	0.293
	2nd Station	0.053	0.052	0.046	0.039
	3rd Station	0.025	0.024	0.022	0.019
Public transport	1st Station	0.286	0.283	0.270	0.255
	2nd Station	0.068	0.067	0.065	0.063
	3rd Station	0.022	0.022	0.021	0.020
Bicycle	1st Station	0.206	0.205	0.199	0.191
	2nd Station	0.031	0.031	0.030	0.029
	3rd Station	0.008	0.008	0.008	0.007
Walking	1st Station	0.122	0.122	0.120	0.117
	2nd Station	0.016	0.016	0.016	0.015
	3rd Station	0.003	0.003	0.003	0.003

In Table 6.6 we also see that the elasticities of the stations accessed by motorized access modes are higher than the corresponding elasticities of stations accessed by bike and walking. This is because the area coverage for which the non-motorized access modes can be used is quite limited. At the branch level, the elasticity of the RSQI on the choice of departure station and access mode follows a reverse pattern as we go from the 1st station to the 3rd station. For each of the access modes, the 1st station has the highest elasticity. Because of the counteracting forces, the resulting total elasticity of the RSQI is rather diffuse in pattern across the three stations accessed by all access modes. In general, the RSQI has a higher elasticity for the 2nd station, with the exception of stations accessed by walking.

6.6.2 Effect of distance

The average number of cars in the postcode areas is 0.402 cars per person. Based on this rate, the utility level of the access modes are plotted in Figure 6.3. Distance is given on the x-axis. All utility curves are downward-sloping, showing the decline in the utility as the distance increases. At any point along the distance range, the utility of one access mode is dominant. We can safely say the access mode corresponding to the dominating utility curve is the most probable mode of access to the departure station in the interval in which its utility is dominant. The graph indicates that walking is the most probable access mode choice for the distance range of up to 1 kilometre. In the range of distances between 1 km and 5 km the bicycle is the most probable access mode choice. Beyond this point, public transport dominates the car alternative, thus, for longer distances public transport remains the most probable access mode choice.

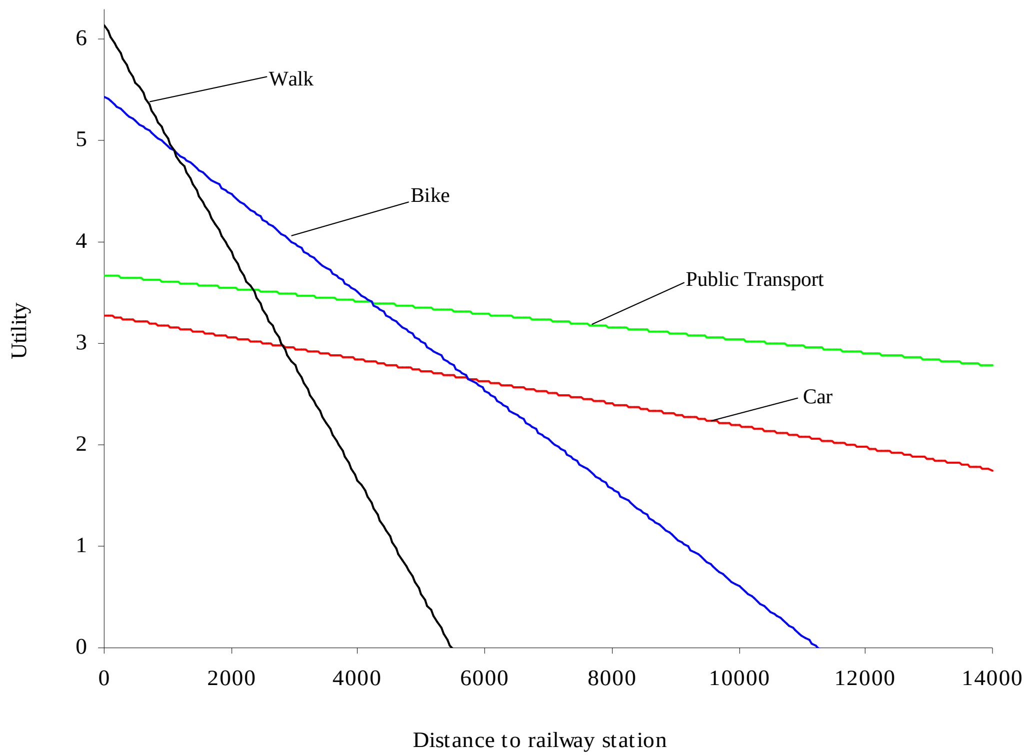


Figure 6.3: The utilities of access modes with respect to distance, for a car ownership level of 0.402 cars per person

6.6.3 The effect of car ownership

The level of car ownership in the postcode area has a positive but insignificant effect on the utility of car access mode. The estimation also shows a positive but highly insignificant effect on bike access mode. However, the estimation shows that the level of car ownership has a negative effect on the choice of public transport access mode. This is in line with our expectation. A higher rate of car ownership in the postcode areas leads to a decline in the choice of public transport. Thus, car mode becomes the most probable access mode of choice, following bicycle for longer access distances, before it is eventually taken over by public transport for further distances. This is the result of competition between public transport and the car. Figure 6.4 plots the utility levels of the access modes setting the car ownership level at 0.60 cars per person, which is above the average car ownership level of 0.402.

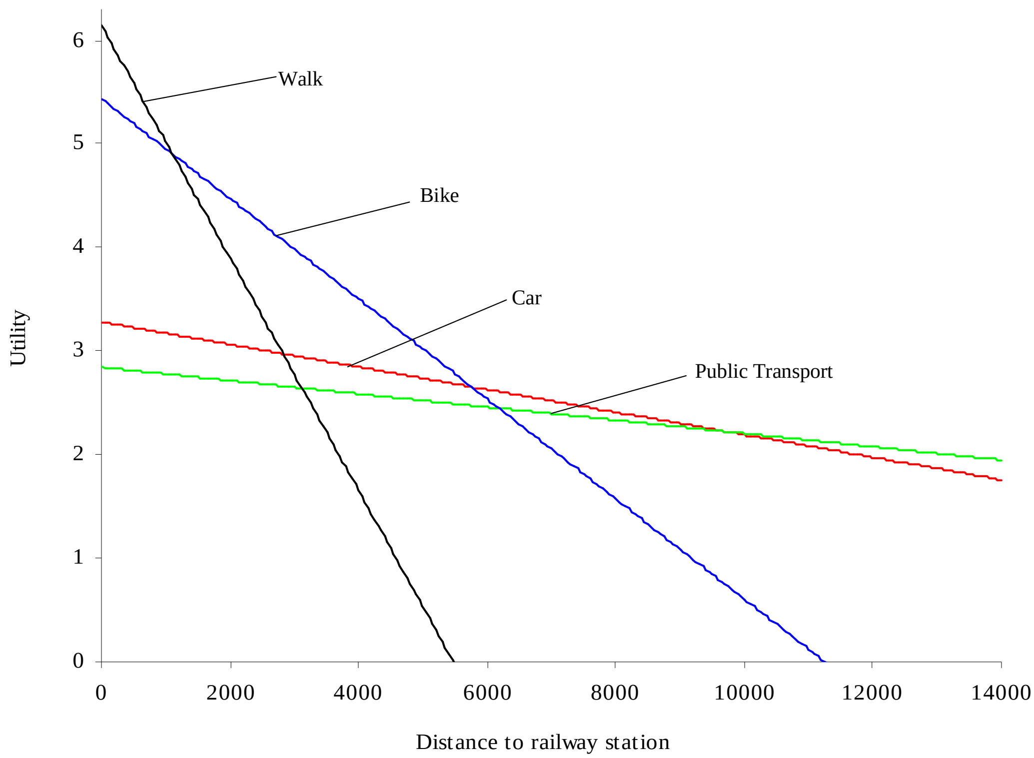


Figure 6.4: The utilities of access modes with respect to distance, for a car ownership level of .60 cars per person

6.7 SUMMARY AND CONCLUSION

This chapter analyses the choice behaviour of Dutch railway travellers concerning the departure railway station and the access mode. Choice is aggregated at the 4-digit postcode area level. For each postcode area a set of four access modes: car, public transport, bicycle, and walking, and a set of three departure railway stations are identified. A rich data set was employed in the analysis. Assuming that these choices are influenced by the assessment of relevant access and station features on the part of the passenger, we distinguish two relevant features for the analysis. The first group includes features related to the ease of accessing the station. In this group we include the distance features that encourage the use of certain access modes to the station. Also included are the levels of car ownership in the postcode area; and the availability of parking area and bike stands at the station level. Features in the second group are related to the rail service delivered at the stations. A comprehensive rail service quality index (RSQI) measure is determined for each station. The RSQI of a station is derived from generalized journey time, distance, and size of destination-railway stations served from the concerned departure station using a doubly-constrained spatial interaction model. It incorporates the frequency of service feature through the waiting time; and service coverage and connectedness though transfer and in-vehicle times. In addition, the importance of the destination stations is accounted for by the size of the destination stations.

A nested logit model was applied to explain choice behaviour concerning departure station and access mode. A nested logit model was estimated based on 1440 postcode areas using a number of access and rail station features. Two structures were analysed. We find that the access mode – departure station choice nesting (from up to down) structure seems more appropriate for the choices than the reverse order nest. The station features used in the estimation include a RSQI and supplementary facilities such as availability of parking space and bicycle stands. The study found the access-mode – departure-station choice nest structure is more appropriate to model the choice process compared with the reverse nest structure. All variables have a significant effect on the choice of access mode and departure station. Distance has a negative effect on the choice of departure station. A steeper effect is observed on the choice of departure stations accessed by the non-motorized modes of walking and bicycle. This implies that they are used on shorter access distances. The level of car ownership has a positive but insignificant effect on the choice of car access mode and a negative effect on the use of public transport. The availability of parking places and bicycle stands has a positive effect on the choice of departure-railway stations accessed by car and bicycle, respectively. Public transport frequency has a positive effect, whereas public transport travel time has a negative effect on the choice of departure stations accessed by public transport. The derived RSQI of a station has a significant and positive effect on the choice of departure stations accessed by all modes. However, the elasticity of the RSQI on the choice of departure station increases as we go from the 1st station with the highest share to the 3rd station with the lowest share for all access mode cases.

Chapter 7

7 The effect of overall railway accessibility on house prices: spatial autocorrelation analyses

7.1 INTRODUCTION

In Chapter 4, we discussed the impact of railway accessibility on residential property prices by means of empirical estimation. A simple hedonic price model was estimated. We found that, after controlling for a number of structural and environmental features of the houses, railway accessibility measured by proximity to a railway station and the service levels provided therein significantly affect the price of dwellings. Because of the spatial nature of the data, it makes sense to explore spatial autocorrelation in the house price data. Dwellings located in the same neighbourhood are generally developed by the same developer, and thus share similar structural characteristics. This leads to the dependence of the price of a certain dwelling on the prices of other dwellings in the neighbourhood. At the same time, dwellings in the same neighbourhood enjoy similar environmental amenities. They are affected by similar policies made at a local administrative level. In the same way, factors that affect house prices such as proximity to the central business district (CBD) or employment area have a similar impact on dwellings in the same neighbourhood. Moreover, some determinants of house price are difficult to measure. All these situations lead to spatially-autocorrelated error components in the hedonic price model. This violates the independence assumption of the error component of the model. Unless properly modelled, the presence of spatial autocorrelation in the data leads to biased estimates.

In this chapter, we apply spatial autocorrelation models in our hedonic price analysis to account for spatial dependence in the house price data. However, the main focus of the analysis remains to determine the effect of railway accessibility on residential property values. This chapter extends the model discussed in Chapter 4 in several directions: 1) it explicitly uses spatial autocorrelation models for the analyses; 2) it utilizes the more comprehensive railway accessibility measure determined in Chapter 6; and 3) it discusses the implications of the HSL (High Speed Line) South at the Amsterdam South Axis for the house prices in the immediate neighbourhood. The terms ‘spatial dependence’ and ‘spatial autocorrelation’ are synonymous. Thus, we will use them inter-changeably.

The chapter is organized as follows. Section 7.2 reviews the empirical literature in the area of the impact of railway accessibility on residential property values. Section 7.3 discusses the modelling approaches in the presence of spatial autocorrelation. This is followed by a discussion on the methodology used for our analysis both in terms of model specification and data used. The model estimation and the discussions are given in Section 7.5. In that section we discuss the projections of the HSL South at the Amsterdam South Axis and their implications for house prices. The chapter ends with a conclusion.

7.2 LITERATURE REVIEW

A large body of the literature on the impact of railway accessibility on residential property values was reviewed in Chapters 2 and 4 of this thesis. In this section we review the literature in the area which involves the use of spatial autocorrelation models. The use of spatial autocorrelation models in real estate price studies is growing. However, the number of studies on the impact of railway accessibility on real estate prices which address the spatial autocorrelation in the hedonic price model remains limited. As far as our search is concerned, spatial autocorrelation analysis has been applied in two studies in the area. The first study by Haider and Miller (2000) analyses the effect of transport infrastructure on residential real estate values. Using an autoregressive spatial hedonic price model, they found that proximity to transport infrastructure explained by proximity to highway and public transport has a significant effect on house prices. They found that dwellings located within 1.5 kilometres of a subway station sell for about 2% more than dwellings located outside this range. In a recent study, Armstrong and Rodríguez (2006) analyse the local and regional accessibility benefits of commuter rail services in Eastern Massachusetts on residential houses, using a spatial hedonic price model. They found that the benefits of railway accessibility are capitalized in house prices. Houses in municipalities which have commuter railway services are about 10% higher in value than houses in municipalities without a commuter rail service. At the same time, their analysis shows that houses located within ½ mile from the station have values that are about 10% higher than houses outside this range. In addition, the negative noise effect associated with a commuter railway measured by the perpendicular distance to the rail line was found to be significant.

7.3 MODELLING SPATIAL DEPENDENCE IN REAL ESTATE PRICE MODELS

The first step in spatial analysis is to trace the presence of spatial dependence in the data. The presence of spatial dependence is generally traced by tests such as Moran's *I* and Geary's C (Moran 1948; Geary 1954; Cliff and Ord 1973). These tests are global in nature, in that they consider the overall data and return a single value that summarizes the dependence status in the data. In contrast, local dependency tests such as LISA Gi and Gi^* statistics (Ord and Getis 1995), LISA statistics (Anselin 1995), local Moran, etc. provide tests for spatial dependence at the local level. However, in this chapter we only apply the global Moran's *I* test for spatial dependence.

Once spatial autocorrelation is traced, the next step is to devise ways to specify the model in a way that incorporates the spatial dependency. The ways of specifying the hedonic price model are dictated by the form of spatial dependence. Two forms of dependence can be distinguished. The first form of spatial dependence is called the lag dependence (or structural dependence) where there is a two-way dependency between the prices of neighbouring residential houses. This implies that the price level of a particular residential house is affected by price levels of other residential houses in the neighbourhood. Ignoring the effect of the price of neighbouring houses in the hedonic price analysis leads to biased and inefficient estimates. To correct for the problem arising from this type of spatial dependence requires the re-specification of the deterministic part of the hedonic price model. This is generally done in two ways. The first and most popular way is to include the dependent variable on the right-hand side by means of a spatial weight matrix (Anselin 1988a). Another approach in specifying the deterministic part of the model involves the inclusion of spatially-lagged independent variable(s) of the neighbouring houses which exhibit spatial autocorrelation (Florax and Folmer 1992). This approach is aimed at avoiding heteroscedasticity in the model. Moreover, Tse (2002) specifies the spatial dependency in the data through the constant term. Generally the spatial lag models are aimed at capturing the spillover effects of neighbouring houses. Thus, in the house price analysis, the coefficients of the independent variables are interpreted as the way they produce a direct and an indirect effect through the effect on neighbouring houses prices.

The second form of spatial dependence is called error dependence. It occurs when the error components of the hedonic price model for neighbouring houses show dependence on each other. In the spatial models this is corrected by specifying the stochastic component of the hedonic price model.

There are other types of models which deal with the spatial dependency. They are called geo-statistical models. They follow the 'kriging' approach for modelling spatial dependency. Geostatistical models have been applied in several real estate data analyses (Dubin 1992; Basu and Thibodeau 1998; Gillen et al. 2001). The kriging approach involves the direct estimation of the variance-covariance matrix by using correlogram or variogram functions. The correlogram and variogram functions can be a function of distance generally known as 'isotropic' (Dubin 1992; Basu and Thibodeau 1998) or a function of both distance and direction in which case it is called 'anisotropic' (Gillen et al. 2001). Based on the estimated variance-covariance matrix, estimation of the regression model is given by estimated generalized least squares (EGLS). The kriging approach mainly focusses on prediction accuracy.

7.4 DATA AND METHODOLOGY

7.4.1 Data

The main data source for the estimation is the database on house sales transactions of the Dutch Brokers Association (NVM). From the 17-year data period analysed in Chapter 4, our analysis in this chapter is based on house sales transactions in the year 2000 in six municipalities: namely, Amsterdam, Rotterdam, The Hague, Haarlem, Almere and Hilversum. These municipalities represent the regions which had the highest number of house sales in that year. In addition, they are mostly located in the Randstad region of the Netherlands. This approach aims to minimize the level of heterogeneity between regions. A further limiting factor in the selection was the computational capacity of the computer. Because of the excessive demand of spatial computation for computer memory, our selection was limited to these six municipalities. A total of 13,058 observations are used in the analysis. In the real estate literature, it is generally assumed that house prices are affected by three categories of features: structural, accessibility, and environmental features. Our data set contained several features in each category. We use approximately the same set of explanatory variables that were used in Chapter 4, with some notable exceptions related to accessibility in general, and railway accessibility in particular.

1. Structural features.

The surface area and number of rooms are some of the important structural features that determine the house prices. In addition, our analysis includes features of the house such as its age, number of bathrooms, monument status, the presence of a garden, garage, gas heater, and open fireplace. The data on the structural features are available in the main data set of house sales transaction records by NVM.

2. Accessibility features

a. Railway accessibility

As we discussed in detail in Chapter 6, railway accessibility has two components: local and regional. The local railway accessibility component of overall railway accessibility explains the ease of reaching the railway station. Accessing the railway station can be done by different modes of transport. Thus, any two given areas which are located in a similar distance range from the railway station can have different local railway accessibility on account of the status of the access modes in these areas. An area which has a high-frequency of public transport connections to the railway station has higher local railway accessibility compared with an area with a low frequency. A similar situation pertains to car ownership. In poor neighbourhoods, where car ownership is low, keeping the status of other modes constant, local railway accessibility there is low compared with neighbourhoods which have high car ownership levels.

The regional accessibility of railway stations is explained by the level of accessibility a station provides to regional destinations. This factor was dealt with extensively in Chapter 5. In this thesis we assume the regional accessibility of a railway station depends on the level of rail service that a railway station provides to other destinations in the railway network. The overall level of service of a particular railway station given by the aggregate sum of rail service across all destinations is a function of the generalized journey time, the generalized journey time to distance ratio, and the importance of the destination station (*see* Chapter 5). In Chapter 4, we showed two considerations of a railway station: the nearest railway station, and the most frequently-chosen railway station. Even though in both considerations we found a similar trend of impact on house prices, they show different magnitudes of the effect. This shows the complexity in accounting for railway accessibility. The choice of a departure station accessed by different modes of transport was analysed in a bid to provide for a comprehensive railway accessibility measure for residential areas (*see* Chapter 6). The comprehensive railway accessibility enjoyed by an area is proportional to the overall utility that households assume in their choice of a departure station (*see* Section 6.3.2). The general railway accessibility levels derived by this procedure are given at the 4-digit postcode area level of aggregation. Thus, general railway accessibility represents the average railway accessibility level in a given postcode area. Houses within the same postcode area are assumed to have the same level of general railway accessibility.

The nuisance associated with railways is accounted for by the measure of perpendicular distance to the railway line. The noise effect is, however, expected to be limited to short distances. Thus, we compare the effect of proximity to the railway line in different distance segments. Two immediate distance segments: namely, within 250 metres, and from 250-500 metres, are compared with distances beyond 500 metres from the railway line. Negative effects are expected, with higher magnitudes for the most immediate segments.

b) Highway accessibility

In Chapter 2, we found that highway accessibility presents an important competition to railway accessibility. It is shown that the exclusion of highway accessibility from the hedonic price estimation results in over-estimation of the impact of railway accessibility. Thus, inclusion of highway accessibility in the estimation procedure is expected to help in singling out the relevant effect of railway accessibility on house prices. Based on the highway network map of the Netherlands for the year 2000, we computed the distance from each dwelling unit to the nearest highway entry/exit point in order to take into account the highway accessibility feature. Moreover, a perpendicular distance measure to the highway is used to account for the nuisance effect of highways on the prices of residential units. Similar to the nuisance effect of railway lines, we compare two immediate distance segments with a reference category which is given by distances beyond 500 metres from the highway line. The source of this data on the location of highway entry/exit points is the Top10Vector of the topographic service of the Kadaster.

c) Accessibility to employment areas, schools, and hospitals

Proximity to employment areas is expected to be an important factor in determining house prices. Determining the proximity to an employment area is a rather difficult task. In the monocentric city case, all jobs are assumed to concentrate in a central core which is mostly referred as the central business district (CBD). Thus, most studies which try to determine the effect of proximity to an employment area do so by the proximity to the CBD. However, because of the increasing de-concentration of jobs out of the historic CBDs, the usefulness of this approach is limited. In this study we account for the proximity to employment areas by considering a fixed number of jobs. We take this fixed number of jobs to be 100,000. Thus, proximity to jobs is measured by the (weighted) average distance to the 100,000 jobs from the location of the dwelling. The data is available at 100-metres by 100-metres grid level. The source of the data is the Netherlands Institute for Spatial Research (RPB). This value is linked to the house price data through GIS linking. Included in the data set are measures of accessibility to schools and hospitals. These are defined by the Euclidian distance between the residential unit and the nearest school which gives secondary education and the nearest hospital.

3. Environmental Features

We use two groups of environmental features in our analysis. The first group includes the proportion of different land use types in the postcode area where the house is located. About 30 land use types are identified. The list and descriptive statistics of the land use types are given in Table 7.1. The second group of environmental variables pertains to the population demography in the neighbourhood area. Included in our analysis are the household income level and the ratio of non-Western foreigners in the neighbourhood. The data on these features are available at the 4-digit postcode area level. The data on these features are obtained from the Central Bureau of Statistics for the Netherlands (CBS).

Table 7.1: Descriptive statistics of the variables

Description	N	Minimum	Maximum	Mean	Std. Deviation
Dependent variable					
Transaction price (euros)		20,420	2,609,236	202,145	164594.370
Independent variables					
Structural features					
house characteristics					
Surface area (square metres)		1	99,998	263	3,492
Building age (years)		0	405	51.813	38.263
Number of rooms		1	18	4.027	1.678
Number of bathrooms		0	4	1.651	0.857
Presence of garage	1,101			0.084	
Presence of garden	6,082			0.466	
Monument status	173			0.013	
Gas heater	1,959			0.150	
Open fireplace	806			0.062	

Continued

Description	**N**	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Types of houses (dummy variables)					
Simple house (reference group)	220			0.017	
Middle-class house	2,829			0.217	
Upper-class house	1,075			0.082	
Villa	337			0.026	
Country house	11			0.001	
Detached house	37			0.003	
Detached house with patio	10			0.001	
Semi-detached house	22			0.002	
Split-level house	34			0.003	
Ground-floor flat	1,119			0.086	
Upstairs flat	2,289			0.175	
Ground and first-floor flat	72			0.006	
House with porch	881			0.067	
Canal house	46			0.004	
Maisonette	509			0.039	
Care flat	30			0.002	
Flat with lift	1,120			0.086	
Flat without lift	1,364			0.104	
Practice house	58			0.004	
Drive-in house	63			0.005	
Farmhouse	2			0.000	
Apartment	930			0.071	
Accessibility Features					
Railway accessibility (index)		-1.356	2.112	0.908	0.534
Highway accessibility (kilometres)		0.025	8.316	2.194	1.400
Distance to 100,000 jobs		0.637	26.741	7.270	5.578
Distance to school (kilometres)		0	5.805	0.731	0.607
Distance to hospital (kilometres)		0.045	7.955	1.680	0.992
Environmental features					
Nuisance features of railway and highway					
Railway line up to 250 m	1,400			0.107	
Railway line 250 to 500 m	1,802			0.138	
Highway line up to 250 m	1,171			0.090	
Highway line 250 to 500 m	895			0.069	
neighbourhood features					
Income (euros)		7,215	20,908	12,507	2,564
Share of non-Western foreigners		0.016	0.817	0.165	0.130
Land use					
cultivation under glass		0	0.584	0.003	0.033
other agricultural use		0	0.712	0.044	0.097
forest		0	0.696	0.037	0.087
extraction of minerals		0	0.022	0.000	0.001
industrial land		0	0.403	0.024	0.059
service facilities		0	0.733	0.038	0.086
other public facilities		0	0.143	0.007	0.021
socio-cultural facilities		0	0.192	0.030	0.033
railway		0	0.443	0.020	0.032

Continued

Description	**N**	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
asphalted road		0	0.219	0.039	0.030
unpaved road		0	0.009	0.000	0.001
airport		0	0.100	0.000	0.002
park or public garden		0	0.410	0.052	0.068
sports park		0	0.213	0.034	0.050
day trip location		0	0.184	0.005	0.016
allotment gardens		0	0.126	0.010	0.024
accommodation		0	0.051	0.001	0.006
dry natural land		0	0.417	0.016	0.068
wet natural land		0	0.055	0.001	0.004
dumping land		0	0.184	0.000	0.006
wreckage land		0	0.033	0.001	0.004
cemetery		0	0.136	0.006	0.023
construction site (firms)		0	0.289	0.006	0.027
construction site (other)		0	0.488	0.032	0.092
other lands		0	0.104	0.002	0.008
water reservoir		0	0.021	0.000	0.001
water with recreational function		0	0.186	0.002	0.017
other water areas broader than 6 m		0	0.363	0.050	0.065
Municipalities					
Amsterdam (reference group)	3,478			0.266	
Almere	1,471			0.113	
Haarlem	1,582			0.121	
Hilversum	1,016			0.078	
The Hague	3,506			0.268	
Rotterdam	2,005			0.154	

7.4.2 Model specification

(A) Standard hedonic price model.

A non-spatial hedonic price model is used for the baseline estimation. House prices are explained by three categories of features: structural, accessibility, and environmental features. A semi-log specification is used for the hedonic price model. Thus, the coefficients of the variables in the model represent percentage effects. The model includes both dummy and continuous variables. The coefficients of the dummy variables represent the percentage effect of the dummy variable on the house prices as compared with a reference variable in the same category. Some of the continuous variables are given in the natural logarithmic form, so the coefficients of these variables represent elasticity effects:

$$\begin{aligned}\ln(\text{tranPrice}_i) = \alpha + \beta'_{HC} \times HouseChr_i + +\beta'_{HT} \times DHouseType_i \\ + \beta_{railacess} \times Railaccess_i + \beta_{hw} \times Hwayaccess_i + \beta_{jobs} \times \ln(Jobsaccess_i) \\ + \beta_{school} \times Schoolaccess_i + \beta_{hosp} \times hospitalaccess_i + \beta'_{railline} \times Drailline_i \\ + \beta'_{hwline} \times DHwayline_i + \beta'_{Neighb} \times Neighb_i + \beta'_{Region} \times Dregional_i + \varepsilon_i.\end{aligned} \tag{1}$$

where, tranPrice_i represents the transaction price of house i; $HouseChr_i$ is a vector of house characteristics for house i, total number of rooms, number of bathrooms, presence of garage and garden for the house, presence of gas heater and open fireplace, monument status, and age of the building; $DHouseType_i$ is a vector of dummy variables representing the type of house i. 22 types of houses are identified in the analysis. The classification of the houses is given by NVM as part of the sales transaction record; $Railaccess_i$ is the railway accessibility measure of the postcode area at which house i is located. It is expected to have a positive effect on house prices in the postcode area; $Hwayaccess_i$ is the variable for accessibility by highway. It is measured by the distance to the nearest highway entry exit point; $Jobsaccess_i$ is the accessibility of house i to employment areas measured by the average distance from the house to 100,000 jobs. It is given in the logarithmic form, so the coefficient represents an elasticity measure of access to jobs on house prices; $Schoolaccess_i$ is the accessibility of house i to schools. It is measured by the distance to the nearest secondary school; $Hospitalaccess_i$ is the accessibility of house i to a hospital. It is measured by the distance from house i to the nearest hospital. Negative signed coefficients are expected for all accessibility measures except railway accessibility, implying that house prices decrease with distance to the nearest highway entry/exit point, to the employment area, to the nearest school, or to the nearest hospital; $Drailline_i$ is a vector of two dummy variables representing at which distance category the house is located from the railway line. This is expected to account for the noise effect of trains. The railway noise is expected to have a localized effect and thus we compare the effect of noise on two nearby distance ranges against the other distance ranges defined by the model; $DHwayline_i$ is a vector of two dummy variables indicating the location of house i in relation to the perpendicular distance from the nearest highway line. These are expected to capture the nuisance effect of a highway on house prices. Because of the presumed localized effect, the distance ranges we compare are given by the two segments of 250 metres each. These segments are compared with a reference segment lying beyond 500 metres from the nearest highway line closest to the railway line. For both the railway and the highway lines effect, we expect a negative effect on both distance segments, with a higher negative effect on the most immediate segments; $Neighb_i$ is a vector of neighbourhood characteristics including income, ratio of non-Western foreigners and share of land use types in the postcode area in which house i is located. They are all given at the 4-digit postcode level. 28 land use types are identified. The income level of the area is expected to have a positive effect on the price level of the houses in the postcode area. In the estimation, the income level is given in the natural logarithmic form. In contrast, a negative effect is expected for the proportion of non-Western foreigners in the postcode area on house prices; $Dregional_i$ is a vector of dummy variables representing the municipality where the house is located; and lastly, ε_i is the error term. The error components are assumed to be independent and identically distributed (iid). Given this assumption on the error components of the hedonic price model, Equation (1) can be estimated through ordinary least square (OLS) methods.

(B) Spatial hedonic price models

In the literature, several studies have diagnosed spatial dependence in real estate price analysis. This called for the use of spatial models in analysing real estate price data. As we have outlined earlier, modelling spatial dependence can be done in a number of ways. In this chapter we focus on the two most applied methods of modelling spatial autocorrelation: namely, the spatial lag model, and the spatial error model. The general-purpose spatial hedonic price model is given as follows:

$$\begin{aligned}\ln(\text{tranPrice}_i) = {} & \alpha + \rho \times \sum_j w_{ij} \times \ln(\text{tranPrice}_j) + \beta'_{HC} \times HouseChr_i + + \beta'_{HT} \times DHouseType_i \\ & + \beta_{railacess} \times Railaccess_i + \beta_{hw} \times Hwayaccess_i + \beta_{jobs} \times \ln(Jobsaccess_i) \\ & + \beta_{school} \times Schoolaccess_i + \beta_{hosp} \times hospitalaccess_i + \beta'_{railline} \times Drailline_i \\ & + \beta'_{hwline} \times DHwayline_i + \beta'_{Neighb} \times Neighb_i + \beta'_{\mathrm{Re}\,gion} \times Dregional_i \\ & + \lambda \times \sum_j w_{ij} \times \varepsilon_j + u_i .\end{aligned} \quad (2)$$

where, ρ and λ are the spatial lag and spatial error coefficients respectively; ε_j is the error component of house j determined by the standard model through ordinary least square (OLS) estimation; w_{ij} is the weight given to the effect of dwelling unit j on unit i. The descriptions of the weight of effects are given in the next section; and u_i is a white noise error component which is independent and identically distributed $N(0,\sigma^2)$.

The type of the spatial model reduced from Equation 2 above depends on the values that the ρ and λ coefficients assume. We find a standard non-spatial hedonic price model when the two coefficients are equal to 0. If λ is fixed at 0, we find a spatial lag model. On the other hand, fixing ρ to 0 gives a spatial error model. Moreover, we find a higher-order spatial autocorrelation model with both spatial lag and spatial error terms when both coefficients are left to be free. In this chapter, we consider the first three cases. As a baseline model, we estimate Equation 1 which is the result of suppressing ρ and λ to be equal to 0. In addition, the spatial lag and spatial error models are considered by suppressing λ and ρ in Equation 2 to be equal to 0, respectively. The higher-order case is outside the scope of this chapter.

It must be noted that the spatial lag model has additional coefficient interpretation implications for the variables in the model. The total effect of a given variable on the price of a house is given by the direct effect on the house and an indirect effect through the effect on neighbouring houses. The total effect of a variable is then given by the $1/(1-\rho)$ factor of the coefficient estimate associated with the variable. The spatial error model does not, however, interfere with the direct interpretation of the coefficients. It only gives the direct effect of the variables on house prices.

7.4.3 Spatial autocorrelation diagnosis

a) Weights matrix

The elements of the weights matrix in (2) are based on the ‘neighbourness’ status of houses. Houses are considered to be neighbours if they are within 1.5 kilometres from each other. All houses which are located in a radius of 1.5 kilometres centred at a given house are considered neighbours of that house. A first-order of neighbourness is considered. Thus, the neighbourness matrix is a matrix with 0 or 1 elements; where 1 indicates that the pairs of houses (given by the row and column of the matrix) are neighbours, and 0 indicates that they are not neighbours. The weights matrix used in the model estimation is thus derived by a row standardization of the neighourness matrix. This means that each element of the weights matrix is equal to the corresponding value of the neighbourness matrix divided by the row sum of the neighbourness matrix. Thus, each row in the weights matrix adds up to 1. All neighbours of a given house have the same weight.

b) Spatial autocorrelation and specification tests

We use the global Moran's *I* test on the OLS residuals for testing spatial autocorrelation in the data. It gives a weighted correlation coefficient for the residuals. For the row-standardized weights matrix ($\mathbf{W}$), the Moran's *I* statistic is given by ($I = \mathbf{e}'\mathbf{We}/\mathbf{e}'\mathbf{e}$), where $\mathbf{e}$ is a vector of OLS residuals. The Moran's *I* test statistic for our data is equal to 0.092 with a t-statistic of 147. This shows a significant positive correlation between the residuals of neighbouring housing units. Thus, this implies that the OLS estimates are biased.

The choice over which spatial model to use for modelling the spatial dependence in the data is made on the basis of the result of the Lagrange Multiplier tests. They are used to distinguish the spatial model which would be appropriate to model spatial dependence in the data. Two forms of Langrage Multiplier exist for both spatial models. The standard Langrage Multipliers test the significance of spatial dependence that can be captured by the specific spatial model. For example, the standard lag Lagrange Multiplier test traces spatial dependence that can be modelled by the spatial lag model. Similarly, the standard error Lagrange Multiplier test traces spatial dependence that can be modelled by the spatial error model. Thus, the standard LM test for one model ignores the spatial dependence that can be modelled by the other spatial model. The standard LM tests for both types of spatial dependence are given as follows (Burridge 1980; Anselin 1988b):

$$LM_{\text{error}} = [\mathbf{e}'\mathbf{We}/(\mathbf{e}'\mathbf{e}/N)]^2 / [\text{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W})]; \tag{3}$$

$$LM_{\text{lag}} = [\mathbf{e}'\mathbf{Wy}/(\mathbf{e}'\mathbf{e}/N)]^2 / D; \tag{4}$$

with $D = [\mathbf{WX\beta})'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{WX\beta})/\sigma^2] + \text{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W})$.

where, $\mathbf{y}$ is a vector of the dependent variable. In our case it is given by the vector of the log price of houses. $\mathbf{X}$ is the matrix of all the independent variables; and $\boldsymbol{\beta}$ is a vector of the corresponding coefficients.

The second form of the LM tests are known as the robust form of LM tests, and give test results which are robust to ignored spatial dependence of the other form. That is to say, the robust LM test gives test results which are robust to ignored spatial lag dependence, and vice versa (Anselin et al., 1996). The specifications of the robust forms of the LM tests are given, respectively, by:

$$\text{Robust } LM_{\text{error}} = [\mathbf{e}'\mathbf{We}/(\mathbf{e}'\mathbf{e}/N) - T(R\tilde{J}_{\rho.\beta})^{-1}(\mathbf{e}'\mathbf{Wy}/(\mathbf{e}'\mathbf{e}/N))]^2/[T - T^2(R\tilde{J}_{\rho.\beta})^{-1}]; \tag{5}$$

$$\text{Robust } LM_{\text{lag}} = [\mathbf{e}'\mathbf{Wy}/(\mathbf{e}'\mathbf{e}/N) - (\mathbf{e}'\mathbf{We}/(\mathbf{e}'\mathbf{e}/N))]^2/[R\tilde{J}_{\rho.\beta} - T]; \tag{6}$$

with, $(R\tilde{J}_{\rho.\beta})^{-1} = [T - (\mathbf{WX\beta})'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{WX\beta})/(\mathbf{e}'\mathbf{e}/N)]^{-1}$ and $T = \text{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W})$.

All the LM test statistics given above are distributed as χ^2 with one degree of freedom. The test result for the spatial autocorrelation and specification tests are given in Table 7.1. The output is computed using the GeoDa 9.5i software. From the table we can see that there is significant positive spatial autocorrelation in our house price data. The LM tests for both spatial models indicate that both types of spatial models can be used to model the spatial autocorrelation present in the real estate data. However, the significance levels of the tests indicate the spatial error model is more appropriate than the spatial lag model.

Table 7.2: Diagnostics for spatial dependence

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.092	146.7	0.000
Lagrange Multiplier (lag)	1	2331.5	0.000
Robust LM (lag)	1	851.5	0.000
Lagrange Multiplier (error)	1	12029.2	0.000
Robust LM (error)	1	10549.2	0.000

7.5 MODEL ESTIMATION AND DISCUSSION

The estimation results of the three models: standard hedonic price model, spatial lag hedonic model, and spatial error hedonic price model are given in Table 7.2. The discussion in the previous section shows that the estimates of the standard hedonic price model (OLS) are biased because of the spatial dependence detected in the error component of the model. Both spatial models can be used to model the spatial dependence in the data, but the spatial error model seems more appropriate. All estimations are done using Geoda 9.5i software. The standard hedonic price model is estimated by ordinary least squares (OLS), whereas the spatial models are estimated using maximum likelihood. Looking at the likelihood and R^2 values shows that the spatial error model has higher explanatory power than the spatial lag model and, of course, than the standard hedonic price model.

The variable of interest in this study is railway accessibility and the associated nuisance effects. The model estimation of the spatial error model shows that railway accessibility affects house prices positively. A unit increase in the railway accessibility index leads to an increase of house prices by about 4%. What this means has been spelled out in Chapter 6. In the next section we will give a more detailed interpretation for the South Axis. Moreover, railway lines pose localized negative effects on house prices. Keeping other things constant, houses located within 250 metres of the railway line and houses located between 250 metres and 500 metres of the railway line sell for 5% and 2% less compared with houses located beyond 500 metres of the railway line. The sign and relative sizes of the coefficients are expected. On the other hand, the spatial lag model finds an effect as high as 12% on house prices for a unit increase in the railway accessibility index. No significant nuisance effect of the railway line on house prices was found.

The remaining discussion on the output of the estimation is given with reference to the estimation based on the spatial error model. The interpretation of the coefficients depends on the nature of the corresponding variables and the way they are used in the estimation. The coefficients related to dummy variables are interpreted as the percentage effects of the variables on house prices in comparison with a given reference group. The coefficients of continuous variables used without log transformation are interpreted as the percentage effects of the variables on house prices for a unit increase in the value of the variable. Coefficients of continuous variables used in the log transformation represent the corresponding variable elasticities of the house prices. They are interpreted as the percentage effect on house prices as a result of a 1% increase in the corresponding variables.

Generally, the structural features of the houses remain the strongest determinants of house prices. A large part of the price variations are explained by the surface area of the houses (with an elasticity of 0.163) and the number of rooms (with an elasticity of 0.456). Other structural features also have a sizable effect on the houses prices. Houses having a garage, garden, a monument status and open fireplace sell at higher prices than their counterparts without these features. For example, a house with a garage sells at about 17% higher than a house without a garage keeping other things constant. Similarly the difference in the price of a house with and without garden, monument status, and open fireplace are about 8%, 16%, and 3%, higher respectively. The presence of gas heater in a house has a negative effect on the price of a house. The number of bathrooms has a modest effect. For every additional bathroom in the house, the price increases by 1.5%, keeping all other things constant. Significant differences in the prices of different types of houses are also observed. Simple houses are taken as the reference type of houses. Prices of country houses, canal houses, farmhouses, and villas (after controlling for all other related features) are among the highest.

The income level of the neighbourhood, with an elasticity of 0.9 is another strong determinant of house prices. The estimation results show that the propertion of non-Western foreigners in the postcode area has a positive effect on house prices. This is contrary to what is expected because the common conjecture is that new immigrants often find a home in relatively cheap houses. However, the two neighbourhood features tend to be highly correlated. Most of the time, high income neighbourhoods tend to have low rates of non-Western foreigners. Thus, the reverse effect for the rate of non-Western foreigners on house prices can be explained by reasons of multi-collinearity. However, to have a better estimation result for the variable of interest in this chapter: namely, railway accessibility, we keep both variables in the model estimation.

In relation to the other features of accessibility the estimation shows significant effects for the distance to schools and highways. However, highway accessibility as determined by distance to the nearest highway entry/exit point has a positive effect. The price of a dwelling becomes high as the distance to the highway entry/exit point increases. This is the reverse of the expected effect. This could be due to the suburbanization effect. The data used constitute houses in a highly urbanized region of the Netherlands. Because of the already higher congestion factors, prices tend to increase further away from the highway entry exit points. The effect of accessibility to a school has the expected sign. For every 1 kilometre closer to a school, house prices increase by about 3%. The estimation of the spatial error model did not find a significant effect of accessibility to jobs and hospitals.

The regional variation on house prices captured by the municipality dummies show that prices in the other municipalities are lower compared with prices of houses in the Municipality of

Amsterdam. However, prices appear to be significantly lower in the Municipality of Almere and The Hague: there, house prices are 50% and 33% lower, respectively, than prices in the Municipality of Amsterdam.

Table 7.3: Estimation result of hedonic price model on house price

	OLS		Spatial Lag (ML)		Spatial Error (ML)	
Variable	**Coefficient**	**t-value**	**Coefficient**	**t-value**	**Coefficient**	**t-value**
Rho/ Lambda			0.490***	42.963	0.959***	155.848
CONSTANT	2.558***	9.877	-1.719***	-6.669	2.000***	5.900
Log surface area	0.169***	35.716	0.159***	35.731	0.163***	37.533
Building age	0.000	0.858	0.000	-0.870	0.000***	-4.071
Log number of rooms	0.456***	56.644	0.460***	60.912	0.456***	61.093
Number of bathrooms	0.016***	4.600	0.014***	4.240	0.015***	4.533
Presence of garage	0.172***	17.710	0.163***	17.776	0.168***	18.854
Presence of garden	0.090***	11.146	0.090***	11.931	0.084***	11.383
Monument status	0.158***	7.263	0.173***	8.469	0.162***	7.983
Gas heater	-0.141***	-20.345	-0.129***	-19.784	-0.132***	-20.649
Open fireplace	0.042***	4.247	0.029***	3.142	0.033***	3.622
Middle-class house	0.120***	6.364	0.149***	8.400	0.165***	9.526
Upper-class house	0.392***	19.459	0.394***	20.800	0.398***	21.547
Villa	0.570***	22.585	0.583***	24.617	0.604***	26.014
Country house	0.749***	9.194	0.686***	8.959	0.758***	10.204
Detached house	0.464***	9.868	0.473***	10.704	0.516***	12.095
Detached house with patio	0.437***	5.181	0.452***	5.712	0.437***	5.698
Semi-detached house	0.403***	6.843	0.427***	7.734	0.469***	8.737
Split-level house	0.251***	5.178	0.245***	5.388	0.231***	5.205
Ground-floor flat	0.014	0.715	0.053***	2.869	0.041**	2.272
Upstairs flat	-0.020	-1.034	0.010	0.567	-0.011	-0.620
Ground and first-floor flat	0.150***	4.225	0.181***	5.407	0.144***	4.415
House with porch	-0.173***	-8.288	-0.084***	-4.277	-0.080***	-4.119
Canal house	0.631***	14.540	0.678***	16.646	0.655***	16.439
Maisonette	0.028	1.275	0.089***	4.285	0.080***	3.948
Care flat	-0.683***	-13.127	-0.617***	-12.631	-0.567***	-11.757
Flat with lift	-0.063***	-3.027	0.010	0.529	0.007	0.368
Flat without lift	-0.098***	-4.845	-0.037*	-1.922	-0.050***	-2.647
Practice house	0.434***	11.215	0.449***	12.365	0.436***	12.343
Drive-in house	0.116***	3.027	0.170***	4.708	0.160***	4.549
Farm house	0.700***	3.792	0.769***	4.440	0.645***	3.830
Apartment	0.255***	12.411	0.286***	14.804	0.262***	13.870
Railway accessibility (index)	0.082***	9.921	0.062***	7.987	0.038***	2.667
Highway accessibility (kilometres)	0.017***	6.448	0.023***	9.475	0.051***	7.580
Railway line up to 250 m	-0.006	-0.665	-0.012	-1.379	-0.050***	-5.327
Railway line 250 to 500 m	-0.021***	-2.786	0.000	-0.025	-0.017**	-2.210
Highway line up to 250 m	-0.010	-1.072	0.001	0.102	0.034***	2.979
Highway line 250 to 500 m	-0.021**	-2.023	0.004	0.421	0.017	1.568
Log distance to 100,000 jobs	-0.062***	-7.385	-0.070***	-8.881	0.023	1.210
Distance to school (kilometres)	-0.064***	-9.716	-0.049***	-7.958	-0.028***	-3.256
Distance to hospital (kilometres)	-0.007**	-1.962	-0.008**	-2.518	0.008	1.206

Continued

Variable	OLS		Spatial Lag (ML)		Spatial Error (ML)	
	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
Ratio of non-Western foreigners	-0.193***	-5.231	0.191***	5.406	0.432***	8.898
Log income	0.892***	33.428	0.695***	26.999	0.890***	27.016
Cultivation under glass	0.651***	8.637	0.076	1.064	-3.068***	-5.864
Other agricultural use	0.159***	5.020	0.087***	2.924	0.065	1.254
Forest	0.422***	8.928	0.135***	3.018	-0.011	-0.177
Extraction of minerals	4.655**	2.545	2.899*	1.688	0.644	0.138
Industrial land	-0.113**	-2.512	-0.128***	-3.031	-0.032	-0.675
Service facilities	0.197***	5.613	0.152***	4.615	0.102**	2.246
Other public facilities	-0.043	-0.246	-0.321*	-1.950	0.073	0.362
Socio-cultural facilities	0.041	0.466	0.445***	5.305	0.310***	3.211
Railway	-0.006	-0.059	-0.036	-0.407	-0.089	-0.834
Asphalted road	-0.707***	-7.049	-0.276***	-2.917	-0.252**	-2.003
Unpaved road	-38.715***	-12.163	-11.735***	-3.871	-7.777*	-1.934
Airport	-0.097	-0.086	1.554	1.478	3.909***	3.032
Park or public garden	0.248***	5.178	0.191***	4.252	0.492***	8.124
Sports park	-0.439***	-7.475	-0.165***	-2.976	-0.095	-1.267
Day trip location	1.748***	7.823	1.959***	9.323	0.960***	3.427
Allotment gardens	-0.183	-1.424	0.316***	2.619	0.844***	5.203
Accommodation	2.079***	3.452	0.214	0.378	0.480	0.627
Dry natural land	0.387***	7.730	0.117**	2.482	-0.192***	-2.927
Wet natural land	-1.348	-0.592	-5.799***	-2.707	-2.199	-0.712
Dumping land	1.933***	4.771	-0.087	-0.228	-1.212***	-2.976
Wreckage land	-2.485***	-3.829	-1.218**	-1.997	-1.040	-0.934
Cemetery	0.409***	3.654	0.515***	4.897	0.800***	6.476
Construction site (firms)	0.763***	6.382	0.654***	5.830	0.340**	2.405
Construction site (other)	0.209***	4.777	0.303***	7.379	0.443***	6.989
Other lands	0.247	0.643	0.180	0.500	2.034***	3.994
Water reservoir	24.373***	3.786	28.167***	4.656	8.719	1.015
Water with recreational function	-1.784***	-11.172	-0.958***	-6.366	-0.913***	-4.303
Other water areas broader than 6 m	-0.117**	-2.257	-0.044	-0.899	0.100	1.552
Almere	-0.511***	-25.623	-0.335***	-17.466	-0.500***	-2.731
Haarlem	-0.396***	-31.632	-0.181***	-14.394	-0.197	-1.109
Hilversum	-0.407***	-22.419	-0.279***	-16.199	-0.072	-0.344
The Hague	-0.660***	-59.233	-0.388***	-32.445	-0.333**	-2.364
Rotterdam	-0.435***	-33.554	-0.213***	-16.306	-0.213	-1.299
Number of observations		13058		13058		13058
R - square		0.807		0.829		0.838
Log likelihood		-850.035		-331.382		-58.158

* = significant at the 10% level; ** = significant at the 5% level; *** = significant at the 1% level.

7.6 IMPLICATIONS OF HST SOUTH IN AMSTERDAM SOUTH AXIS

The implementation of the HSL South in the Amsterdam South Axis concerns the largest infrastructure-related urban development project in the Netherlands. Near the South Axis, new railway infrastructure has recently been completed. This infrastructure allows for more train services to/from the Amsterdam South Axis Station and can therefore be expected to influence real estate prices in this area. A most notable railway development is the HSL South high-speed railway from Schiphol airport to Rotterdam and further to Belgium. After the HSL South is put into service (foreseen in 2007 or 2008), travel times will be significantly shortened, both from domestic services to Rotterdam and places in the south of the Netherlands; and also to international destinations (Brussels and Paris).

The question whether the Amsterdam South Axis Station will accommodate high-speed train services is still uncertain. In the coming years, capacity restrictions at the station will make it impossible to have high-speed train services. Therefore, in the early years, these services will stop at Amsterdam Central Station instead. In the long term both Amsterdam South Axis Station and Amsterdam Central Station are options. In our analysis, we assume that all high-speed train services towards the South will use Amsterdam South Axis Station.

The domestic connections account for the largest part of the train services that will make use of the new high-speed railway. According to the projected schedule (High Speed Alliance 2006) of all 96 trains leaving Amsterdam per day only one-third goes to Belgium, and half of these continues to Paris. With this schedule, travel times between Amsterdam and Rotterdam will decrease from 53 minutes to 30 minutes and between Amsterdam and Breda in the South of the Netherlands from 1 hour and 42 minutes to 54 minutes. This reduction of travel times will have a large positive impact on the RSQI of the South Axis Station.

Besides the high-speed train services, the Amsterdam South Axis Station will receive several other improvements in railway services. Early in 2006 a new direct intercity connection to Utrecht and Eindhoven was introduced, which forms an important link for the South Axis. In addition to this, the new railway schedule that is proposed by the Dutch national railway company (NS 2006) implies a further increase of train frequencies at the South Axis Station. This mainly concerns regional train services. The operation of the HSL in the Amsterdam South Axis is expected to lead to an increase in the regional railway accessibility as provided by the South Axis station. This in return is expected to affect house prices positively. The projections discussed here only consider the foreseen changes related to the HSL South. Other changes in the regular rail operations are not considered.

The South Axis Station is included in the departure station set of several postcode areas. For an assessment of the implications of the HSL South operation, we will concentrate on the postcode areas which are expected to benefit the most. The South Axis Station is located on the boundary of two postcode areas in Amsterdam; these are 1077 and 1082. The implementation of the HSL South at the station will lead to an increase in the general railway accessibility level of these areas. Under the current settings, the model estimation for general railway accessibility discussed in Chapter 6 predicts that the operation of the HSL South in South Axis station will lead to an increase of the general railway accessibility measure by 0.7 (from 1.17 to 1.87) and 0.72 (1.20 to 1.92) for postcode areas 1077 and 1082, respectively. For this level of improvement in the railway accessibility for the immediate postcode areas, the spatial error model predicts an increase of about 3% on house prices in these postcode areas.

7.7 CONCLUSION

In this chapter we have analysed the effect of railway accessibility on house prices. The analysis is based on house sales transactions in six municipalities of the Netherlands for the year 2000. Using spatial dependence test results a positive spatial dependence is diagnosed. Thus, to correct the effect of the spatial dependence in the data, spatial autocorrelation models are used for estimations. The use of a spatial autocorrelation model considerably improves the estimation result in comparison with estimation results of a non-spatial hedonic price model. However, even though both the spatial lag and spatial error models were found to be significant to model the spatial dependence in the house price data, the spatial error model has more explanatory power than the spatial lag model.

Controlling for several structural and environmental features, we found a positive effect for the general railway accessibility index on house prices. Using the spatial error model, we found that a unit increase of the index in a postcode area leads to about a 4% increase in house prices in the postcode area. The railway accessibility index of an area could increase by improving the access and rail service quality features of the railway stations used for departure by household in the area. The projection concerning the effect of the foreseen HSL South operation in the Amsterdam South Axis shows the impact size of changes on the components of the railway accessibility index (RSQI). The HSL-South will lead to the reduction in rail trip travel times, and this in turn will increase the rail service quality of the station. The model estimations in Chapter 5 predicts that the operation of the HSL South on four stations (Amsterdam South, Schiphol Airport, Rotterdam, and Breda) with the foreseen timetable will increase the rail service quality of the station by 80%. This leads to an increase of about 0.7 in the general railway accessibility index of the two postcode areas close to the Amsterdam South Axis station which is currently the dominant departure station for those areas. This increase implies about a 3% increase in house prices in these areas. A similar increase in price can be expected for houses located in other postcode areas which use the station for departure or will start to use the station because of the improvement in rail service, though the size of the effect may be smaller than that for the immediate neighbourhood of the station.

Chapter 8

8 The effect of railway stations on commercial property values: a spatial autocorrelation model

8.1 INTRODUCTION

Railway station surroundings are sometimes known as the “shop window” of a city because they serve as places where people can see what the community has to offer. Given this special nature of railway-station surroundings, several types of businesses find it attractive to locate themselves around railway stations. They find it attractive because the railway station provides contact opportunities with their customers, with an expected higher sales turnover as a result. In addition, being close to a railway station gives employers access to a potential pool of employees at a reduced cost. This is expected to contribute to the competitive advantage of a firm compared with its counterpart businesses located further from stations. These dual advantages make it possible for commercial entities to be willing to pay a premium on rents to in order to remain close to the railway station. Commercial land rent is, as a result, expected to decline as the distance from the railway station increases.

In discussing the effects of railway stations, it is important to note the distinction between railway stations at the origin and at the destination side of trips, because the departure and destination features of a railway station have different implications for residential and commercial property values. For example, in the decisions undertaken by households for the location of their residence, their decision is likely to be influenced by the assessment of a railway station as a departure station which provides access to an important destination station where they can engage in variety of activities. On the other hand, in the decisions undertaken for the location of their business, business entities mostly assess the value of a station by its trip-attraction quality as a destination station. This will be the point of departure for assessing the value of a station that we follow in this study. Thus, the distance from the nearest railway station to the location of a commercial property under consideration represents the egress part of a rail trip. In the egress part of a trip, visitors or employees mostly rely on walking to get to the location of the commercial property. On the activity end of a railway trip, walking accounts for about 46% of the share of access to and from the railway stations (Rietveld

2000). Thus, the distance range at which the influence of station proximity on commercial land rent is felt is expected to be quite limited.

The aim of this chapter is to analyse the effect of railway accessibility on office rents. Railway accessibility is measured by two features: proximity to stations, and rail service quality. Based on the Dutch office rent market, a hedonic spatial autocorrelation model is estimated. In addition, the chapter discusses the implications of high-speed rail implementation in the South Axis Station in Amsterdam for the rent levels of office space there.

The discussion in this chapter is organized as follows. In Section 8.2 we review the literature on the effect of railway stations on office rent levels. First, we discuss hedonic empirical studies in an international context. Then we briefly discuss the location factors for offices in the Netherlands. This will be followed by a review on the effect of high-speed trains. In Section 8.3 we discuss the data and methodological approach. After giving the autocorrelation diagnosis in Section 8.4, we discuss the estimation results of the spatial hedonic price model in Section 8.5. Section 8.6 is devoted to the discussion of the implication of the implementation of HSL South in the South Axis Station for the office rental market. The chapter closes with conclusions in Section 8.7.

8.2 LITERATURE REVIEW

8.2.1 Railway accessibility in hedonic pricing studies in an international context

In the literature, we observe studies which approach the effect of railway stations on real estate from two angles: effects on land use, and effects on property values. In this respect, we review some of the studies on the effect of railway stations on commercial properties from the perspective of both approaches. In one of the earliest studies, Quackenbush et al. (1987) studied the impact of the Red Line in Boston on land use. They found that the largest effect was on commercial properties, with only a slight effect on residential properties. In addition, Weinstein and Clower (1999) indicated that on the announcement of Dallas Area Rapid Transit (DART), the occupancy rate of commercial properties within ¼ mile of the stations increased on average by 5%. A number of different findings on the effect of railway stations on commercial property values are found in the literature. The study on the effect of proximity to a metro station on commercial property values in Washington D.C. was one of the early studies in this regard (Damm et al. 1980). The study found that the values of commercial properties decline with distance. Proximity to a metro station results in a steeper effect on commercial property values as compared with the effect on residential property values. The elasticity of proximity to the railway station on property values was around 4 times higher for commercial properties than for residential properties. This shows that, in the immediate neighbourhood the premium of closeness to a station is greater on commercial properties. Commercial land value premiums were also found by Fejarang (1994). He found that commercial space in Los Angeles city that is located within ½ a mile of a rail transit station had an additional $31 increase in mean sales price per square foot over comparable parcels outside the corridor. In addition to the land use changes as a result of the announcement of the opening of DART, Weinstein and Clower (1999) observed an increase in the rent of three classes of offices within ¼ mile of a station ranging from 20.9% to 47.4% compared with the same kind of offices outside that range. Similarly the study done by the US Federal Transit Administration (FTA) indicates the price per square foot of commercial space decreases by about $2.3 for every 1000 feet further from a railway station. This value accounts for approximately 2% of the value (FTA 2002). Furthermore, Nelson (1998) found that the price per square metre in Atlanta decreases by $75 for every metre further away from a transit station. In an effort to present background information against the law suit brought by private property owners in Santa Clara County, claiming a burden due to the existence of light rail transit (LRT), Weinberger (2001) tested several hedonic price models on the rental rates of commercial property. The finding reveals that the results of commercial properties within ¼ mile of the station are 10% higher than rents of commercial properties beyond ¾ mile of a light rail station. When controlling for highway access, the rail proximity benefit was maintained, and it was shown that highway coverage in the county is so dense that there are no particular locational advantages associated with highway coverage. A similar study was also done by Cervero and Duncan (2001) in the same County. They found that commercial properties within ¼ mile of a light railway station were sold at prices 23% higher than commercial properties outside this range. The capitalization is even more pronounced in the case of proximity to a commuter railway station. Commercial properties within ¼ mile of the commuter railway station sell at prices more than 120% above commercial properties outside this range. But, contrary to the above positive effects of proximity to a railway station, Cervero (2001, 2002) found a mixed effect of proximity to a railway station on commercial property values. The study in Los Angeles County shows that for commercial properties located within ¼ mile of a station, the impact of different stations ranges from a negative effect as big as 30% to a positive effect of 16% compared with the values of properties outside the ¼ mile range. Similarly, the study in San Diego County reveals that the impact of proximity to a railway station within ¼ mile on commercial property values ranges from a negative effect of 10% to a positive effect as big as 90%. Landis et al. (1995) found no premium for commercial land. However, the inability to find a positive impact is attributed to a data and methodological problem rather than to of the lack of a real value premium.

8.2.2 Location factors in the Netherlands

The price of real estate depends on the attractiveness for decision makers to choose is as a site for their activities. Besides hedonic price models different types of studies can be distinguished that aim to identify and rate the location factors that underlie the attractiveness of locations. Most common studies are surveys of economic-geography in which an importance ranking of location factors is provided, based on questionnaires. These surveys typically take account of a large number of location factors and can therefore provide a good overview of the field. For the Netherlands a great number of studies of this type have been conducted (see, e.g., Pellenbarg 1985; Jansen and Hanemaayer 1991; Sloterdijk and van Steen 1994; and, for an international study that includes the Netherlands, see Healey & Baker 1996). Besides questionnaires, other more advanced quantitative methods can also be used. Examples in the Netherlands include stated choice studies (e.g. Rietveld and Bruinsma 1998) and advanced rating studies (Berkhout and Hop 2002). The number of location factors in these studies is normally smaller.

Little uniformity exists in the specification of location factors. Nevertheless, when examining these studies, four general categories of location factors can be distinguished:

1. Accessibility-related: includes proximity to actors or infrastructure, and the availability of personnel;
2. Properties of the building: includes availability, representativeness, possibilities for expansion;
3. Properties of the surroundings: includes type of environment, representativeness;
4. Other regional factors: includes working mentality, quality of life, investment subsidies.

The analysis in this chapter looks at the effect of the first three factors on office rent levels in the Netherlands. However, the discussions focus on the first of these factors: accessibility-related factors which are the main interest of this chapter. Accessibility-related aspects are among the most important location factors. Aspects of accessibility in location studies are, on the one hand, the connectivity to the network (access distance or travel time to a network node, or the level-of-service of this node) and, on the other hand, the potential accessibility (relates to the possibility to reach destinations, e.g. as a result of the availability of potential personnel). In general, accessibility by car is seen to be more important than accessibility by public transport (e.g. Jansen and Hanemaayer 1991). However, accessibility is typically not analysed in much detail. In addition, the link to the price of real estate is also weak. In this chapter the effect of both railway and highway accessibility on office rent levels is studied. The biggest volume of literature on real estate value in relation to railway stations relates to light, heavy, and commuter railway stations. In this chapter we are interested in the effect of commuter railway stations in the Dutch railway network. The estimated model is used to predict the implications of high-speed railway implementation.

8.2.3 High-speed rail and location attractiveness

For the South Axis Station in Amsterdam an important accessibility feature is the possibility to have high-speed train services via the nearby new HSL (High Speed Line) South. Domestic services to Schiphol Airport, Rotterdam and Breda as well as international services to Antwerp, Brussels and Paris are among the possibilities. High-speed rail connections can improve long-distance accessibility and therefore also location attractiveness and real estate prices. The extent to which the proximity of high-speed rail can raise real estate prices is still uncertain. No studies that analyse the impact of high-speed rail on real estate prices are known to the authors. Indications can be derived from various other types of studies on the spatial effects of high-speed rail.

In countries with high-speed railway lines, empirical studies have focussed on the spatial-economic effects of high-speed rail, both at an interregional and intraregional scale. On an interregional scale, studies in Japan have showed the existence of a statistical relationship between the presence of a Shinkansen station and regional growth. Hirota (1984, as referred to by Brotchie 1991) found a positive correlation between the presence of a Shinkansen station in a city and growth indices for several economic sectors and for population, even though the cities with a Shinkansen station have had lower growth rates on average than other cities before the Shinkansen was opened. Nakamura and Ueda (1989, as referred to by Brotchie 1991) found a similar result when comparing regions with and without a Shinkansen station, which was further enhanced when the presence of an expressway was also taken into account. Although these studies provide useful information, they are not conclusive on the causality of the relationship found. Besides the impact of Shinkansen on regional growth, there is also the possibility that the government decision to link a city to the Shinkansen was taken in anticipation of the expected growth of the city.

A number of descriptive research studies on firm relocations, using entrepreneurial surveys, have studied the effect of high-speed rail on the urban or intraregional level. Entrepreneurial surveys can shed light on the motives of location decisions and the role of high-speed rail. This type of research has been carried out mainly in France, and includes studies reported by Bonnafous (1987), Sands (1993) and Mannone (1997). As a general conclusion for France (Haynes, 1997), the TGV was of minor importance for the location decisions of most firms. In most cases high-speed railway accessibility is just one of a series of factors that influence location decisions. Industrial firms are particularly constrained in their location choice by other factors. In a sample of entrepreneurs located near the Lyon Part-Dieu high-speed railway station, Mannone (1997) found only about one-third of the respondents indicated that the high-speed train services had been a predominant factor in their location choice. For the case of Grenoble, Mannone (1997) suggests image effects to be relevant, as is also mentioned by Sands (1993) for the city of Nantes. However, the importance of image effects on location attractiveness is difficult to assess from these studies.

From these studies it can be concluded that high-speed rail connections can influence real estate prices, but they are not expected to be dominant. Related aspects such as station area redevelopments and improved regional accessibility can be at least as important as the high-speed trains themselves.

8.3 DATA AND METHODOLOGY

8.3.1 Data source

The main sources of data for the estimations in this chapter are the recoded office rent contracts from Zadelhoff DTZ. It extends over a period of 23 years from 1983 to 2005. Geographically it covers all provinces of the Netherlands. The data set includes information on the rent per square metre of office floor space, building status, type of rent contract, and category of business. Five types of building status are identified. These are: first-user buildings; second-user buildings; buildings under construction; buildings under renovation; and land yet to be developed. There are three types of rent contracts: namely, first-hand new rent contracts; rent extension contracts; and subleases. In addition, the data set identifies the type of business the user conducts. One may claim that the data relate to the user and not to the building itself. However, these data can be used as a proxy for the nature of the building, because the nature of the building required can differ according to the business orientation of the user. For instance, the type of building required by a banking or insurance company is generally different from that required by a transportation and storage business. A number of other variables are also included. To account for the environmental features, we include the share of different land use types in the postcode area. Because the data includes rent contracts for a long period, year dummies are included to capture the temporal change in the rent levels related to inflation and real value changes.

Two types of accessibility variables are included. Railway accessibility is measured by the proximity of the office location to the nearest railway station. Accessibility by road is measured by the distance to the nearest highway entry/exit point from the location of the office. In order to compute these distance measures, the stations, railway line, highway entry/exit points and office location had to be geo-coded. Coding was done at a detailed address level, because the office rent is generally expected to be sensitive to distance and, according to the literature, the range of distance at which the rent of commercial properties responds to proximity is rather limited. To account for the effect of business locations' opportunities for interaction with customers and employees on the rent level, we introduce a derived opportunity index for the business locations. The statistical tests over the different opportunity (accessibility) indicators made by Song (1996) indicate that gravity type opportunity measures generally perform better than other measures. We define the opportunity index as the cumulative population of all postcode areas in the country weighted by the inverse of distance from the office location to the centroid of the postcode area:

$$OpportunityIndex_i = \sum_{j=1}^{R} \frac{Population_j}{d_{ij}} . \tag{1}$$

where, d_{ij} is the Euclidean distance between the location of office i and the centroid of postcode area j. The opportunity index used in the empirical estimation of this chapter is based on all the 4004 postcode areas that cover the whole Netherlands. Distance is measured in metres.

Railway accessibility is explained by two variables: a measure of the rail service quality at the nearest railway station, and the distance to the nearest railway station. We use a derived index for the measure of rail service quality. As pointed out earlier, the destination point of view of the station is of more relevance to explain the office rent levels. The derivation of this measure is discussed in Chapter 5. The descriptive statistics of the data used in the estimation are given in Table 8AI.1 in Appendix 8AI.

8.3.2 Methodology: Econometric models

(A) Standard hedonic price models

The analysis in this chapter is based on hedonic pricing model estimation. The variables of interest are related to accessibility in general, and railway accessibility in particular. There are two types of variables related to railways: distance to the nearest railway station, and the rail service quality index (RSQI) at the nearest station. These variables are expected to capture the effect of railway accessibility. In addition, the model includes accessibility to the highway. It is measured by the distance to the nearest highway entry/exit point. Furthermore, the model includes access to potential employees and business customers by assuming a radial access to the office location. A semi-logarithmic hedonic model is specified. The dependent variable is given in the natural logarithmic form; thus, the values of the coefficients represent percentage change. The specifications of the standard models used in the estimations are given by Equations 2 and 3. Distances from the offices to the nearest railway station are divided into 6 distance categories, where distances above 4 kilometres are taken as the reference group. In the first model, distance and RSQI are included separately. In the second model, however, a cross product of distance and RSQI is included with the aim of observing the effect of service quality on office rent levels at different distance classes. The respective base models have the following form:

$$\begin{aligned}\ln(\mathrm{rent}_i) = \alpha &+ \beta'_{BS} \times DBuildStatus_i + \beta'_{CT} \times DContrType_i + \beta'_{BT} \times DBusType_i \\ &+ \beta'_{dc} \times Distcategrail_i + \beta_{RSQIdest} \times RSQIdest + \beta_{hw} \times \ln hwdist_i \\ &+ \beta_{opport} \times OpportunityIndex_i + \beta'_{Neighb} \times DNeighb_i + \beta'_{time} \times Dtime_i + \varepsilon_i;\end{aligned} \tag{2}$$

$$\begin{aligned}\ln(rent_i) = \alpha + \beta'_{BS}\times DBuildStatus_i + \beta'_{CT}\times DContrType_i + \beta'_{BT}\times DBusType_i \\ + \beta'_{dc}\times Distcategrail_i \otimes RSQIdest + \beta_{hw}\times \ln hwdist_i \\ + \beta_{opport}\times OpportunityIndex_i + \beta'_{Neighb}\times Neighb_i + \beta'_{time}\times Dtime_i + \varepsilon_i .\end{aligned} \tag{3}$$

where, $rent_i$ is the rent per square meter of space for office i, $DBuildStatus_i$ is the building status of office i; $DConstType_i$ is the rent contract type of office i; $DBusType_i$ is the business type for office i; $Distcategrail_i$ is the category for the distance at which office i is located from the railway line. A positive sign is expected for coefficients for the distance categories with the highest effect in close proximity to the nearest station and decreasing outwards; $RSQIdest_i$ is the RSQI of the nearest station for office i. A positive effect is expected, showing that an increase in the RSQI of the nearest railway station to the office location leads to higher rents; $hwdist_i$ is the distance between office i and the nearest highway entry/exit point. We expect the office rents to decline as the distance to the nearest highway entry/exit point increases; $OpportunityIndex_i$ is the opportunity index for office i (defined by Equation 1). The greater the opportunity for interaction with potential customers and employees, the higher the office rent level; $Neighb_i$ is the land use feature for the postcode areas in which office i is located; $Dtime_i$ is a time dummy variable representing the year when the rent contract took place; and lastly, ε_i is the *iid* error term. The definition and the descriptive statistics of the variables are given in Table 8AI.1 in the Appendix 8AI.

(B) Spatial hedonic price models

The standard hedonic price models discussed above assume that rent levels of offices in the sample are independent from each other. However, as the law generally referred to as Tobler's first law of geography states "everything is related to everything else, but near things are more related than distant things" (Tobler 1970), it is not impossible that the assumption is violated. This is because offices in the same area tend to share similar physical, environmental and accessibility features. This results in spatially-correlated rent levels. At the same time, location-related characteristics are generally difficult to observe and quantify, and the omission of variables from the hedonic price model results in spatially-correlated error terms. The violation in the assumption of independence of the error term leads to inefficiency in the parameters estimate by ordinary least squares (OLS) methods. In the literature, two ways of dealing with the spatial dependence are proposed. The first approach includes the weighted average of neighbouring office rents. The second approach involves modelling the structure of the error term of the standard model. The general cases of spatial hedonic price models corresponding to the earlier discussed models are given below:

$$\begin{aligned} \ln(\mathrm{rent}_i) = {} & \rho\sum_j (w_{ij} \times \ln(rent_j)) + \alpha + \beta'_{BS}\times DBuildStatus_i + \beta'_{CT}\times DContrType_i \\ & + \beta'_{BT}\times DBusType_i + \beta'_{dc}\times Distcategrail_i + \beta_{RSQIdest} \times RSQidest \\ & + \beta_{hw} \times \ln hwdist_i + \beta_{opport} \times OpportunityIndex_i + \beta'_{Neighb}\times DNeighb_i \\ & + \beta'_{time}\times Dtime_i + \lambda\sum_j (w_{ij} \times \varepsilon_i) + u_i; \end{aligned} \quad (4)$$

$$\begin{aligned} \ln(\mathrm{rent}_i) = {} & \rho\sum_j (w_{ij} \times \ln(rent_j)) + \alpha + \beta'_{BS}\times DBuildStatus_i + \beta'_{CT}\times DContrType_i \\ & + \beta'_{BT}\times DBusType_i + \beta'_{dc}\times Distcategrail_i \otimes RSQIdest + \beta_{hw} \times \ln hwdist_i \\ & + \beta_{opport} \times OpportunityIndex_i + \beta'_{Neighb}\times Neighb_i + \beta'_{time}\times Dtime_i + \lambda\sum_j (w_{ij} \times \varepsilon_i) + u_i. \end{aligned} \quad (5)$$

where, ρ and λ are the weighted lag and error coefficients; w_{ij} is an indicator of the neighbourness of office j to office i in the row standardized weights matrix; ε_i is the residual of the OLS estimate for office i; and u is white noise error term $(u \sim N(0,1))$. If $\lambda = 0$, the model reduces to the spatial lag model. In this case, the office rent level is dependent on the weighted average rent of the neighbouring offices. But, if $\rho = 0$, the model reduces to the spatial error model. If both coefficients are different from 0, we get a higher-order spatial specification that involves both spatial lag and spatial error models. In this chapter, the estimation considerations will be limited to the case where either one of the two coefficients is 0.

8.4 DIAGNOSIS FOR SPATIAL AUTOCORRELATION

To assess the spatial dependency in the office rents, we use the Moran's *I* test. A row-standardized weights matrix of neighbourness, based on a 3-kilometres cut-off distance is used, to model the spatial structure of the dependency. By showing the level of spatial dependency on the data, the Moran's *I* test gives an indication of whether the standard (non-spatial) model is misspecified or not. However, the test does not give any information on which spatial model is appropriate for the data. Identifying the appropriate spatial model is based on Lagrange Multiplier tests (Anselin 1995). Table 8.1 gives five Lagrange Multiplier test results. The first two (LM lag and Robust LM lag) are tests on the appropriateness of the spatial lag model. The next two (LM error and Robust LM error) relate to the spatial error model as an alternative model. The last Lagrange Multiplier test relates to a test for a higher order alternative specification that involves both spatial lag and spatial error terms. The specifications of the entire test statistic are given in Appendix 8AI.

The values of the Moran's *I* are positive and significant. This indicates that the error components of the standard model for neighbouring offices are positively correlated – a violation in the independence assumption of the error term. Thus, the ordinary (non-spatial) model estimations result in biased estimates. This calls for the use of a spatial autocorrelation model for the rent data. The choice of the appropriate approach to model the spatial autocorrelation on the data is based on Lagrange Multiplier tests. Two categories of Lagrange Multiplier tests are proposed: a standard and a robust form for each of the modelling approaches as separately. The specifications of the test statistic are given in Appendix 8AI. Both the standard forms of the Lagrange Multiplier tests (LM lag and LM error) are significant, indicating that both spatial lag and spatial error models can be used to model our data. However, of the robust forms, only the robust Lagrange Multiplier test is significant. This indicates that the spatial error model is the preferred model for the data. Using these test results, we apply the spatial error model to model our data.

Table 8.1: Diagnostics for spatial dependence

TEST	MI/DF	VALUE	PROB
1. Separate effect of distance and station's rail service quality index (RSQI)			
Moran's *I* (error)	0.1088	56.75	0.000
Lagrange Multiplier (lag)	1	78.89	0.000
Robust LM (lag)	1	0.48	0.488
Lagrange Multiplier (error)	1	2911.72	0.000
Robust LM (error)	1	2833.31	0.000
Lagrange Multiplier (SARMA)	2	2912.20	0.000
2. Cross-effect of distance and station's rail service quality index (RSQI)			
Moran's *I* (error)	0.1098	57.11	0.000
Lagrange Multiplier (lag)	1	84.47	0.000
Robust LM (lag)	1	0.15	0.698
Lagrange Multiplier (error)	1	2969.62	0.000
Robust LM (error)	1	2885.30	0.000
Lagrange Multiplier (SARMA)	2	2969.77	0.000

8.5 ESTIMATION AND DISCUSSION

The resulting spatial autocorrelation diagnosis discussed above shows that the spatial error model is the preferred model. Thus, we estimated both the ordinary least squares (OLS) and spatial error hedonic price model (SEM) for each of the base models. The spatial autocorrelation models are estimated using Geoda 9.5-i5 software. The estimation results related to the variables of interest are given in Table 8.2. The coefficients of the remaining variables of the models are given in Appendix 8AI. The first set of estimates is based on the model which presents the effect of proximity to railway stations and service quality at the station separately. The second set of estimates is based on the model which treats railway accessibility as a cross-product of distance and the RSQI. A piecewise approach is used for the distance to the nearest railway station. The effect of proximity to the nearest railway station on office rent levels is inferred by reference to the rent levels of offices located beyond 4 kilometres from the nearest railway station. Our discussion will be based on the spatial hedonic error models (SEM). The spatial error parameter in both spatial error models (λ) is equal to 0.71, and highly significant. This shows that the unobservable components of the model for neighbouring offices are positively correlated.

Table 8.2: Estimation results for the effect of accessibility variables on office rent levels

	Separate effect of distance and rail service quality index		Cross-effect of distance and rail service quality index	
Variables	OLS	SEM	OLS	SEM
CONSTANT	2.499*** (23.271)	3.132*** (14.584)	2.487*** (23.452)	3.138*** (14.615)
raildist0_250	0.097*** (6.147)	0.139*** (5.934)	0.169*** (9.826)	0.176*** (8.886)
raildist250_500	0.123*** (8.923)	0.127*** (5.704)	0.221*** (14.712)	0.176*** (9.800)
raildist500_1000	0.072*** (5.758)	0.074*** (3.423)	0.141*** (11.458)	0.087*** (5.445)
raildist1000_2000	0.051*** (4.211)	0.038* (1.836)	0.115*** (10.265)	0.054*** (3.585)
raildist2000_4000	0.035*** (2.844)	0.034* (1.689)	0.091*** (7.409)	0.050*** (3.187)
Rail service quality index (RSQI)	0.144*** (12.749)	0.080*** (5.246)		
Ln(opportunity index)	0.376*** (22.320)	0.267*** (7.437)	0.387*** (23.211)	0.274*** (7.710)
Ln(hwdist)	-0.043*** (-11.581)	-0.044*** (-9.634)	-0.042*** (-11.245)	-0.044*** (-9.652)
Lambda		0.709*** (36.555)		0.707*** (36.252)
Number of observations (N)	**9,357**	**9,357**	**9,357**	**9,357**
R-squared	**0.3602**	**0.4255**	**0.3603**	**0.4257**

* = significant at the 10% level; ** = significant at the 5% level; *** = significant at the 1% level.

8.5.1 Effect of accessibility on office rent

(A) Railway accessibility

From Table 8.2 we can see that the proximity to a railway station has a positive effect on office rent levels. The spatial error model estimation on the separate effect of proximity and rail service quality shows offices within 250 metres of a railway station have a rent of about 14% above that for offices which are beyond 4 kilometres of a railway station. A downward-sloping effect is found: the effect of proximity to a railway station on office rent levels decreases as the distance away from the railway station increases. A statistically weaker positive effect of station proximity is found for offices between 1 km and 4 km compared with offices located beyond 4 km from the nearest railway station. This confirms the assertion that the effect of proximity to the railway station on commercial property is limited to the walking distance range (*see* Chapter 2). A graphical description of the effect of proximity to the railway station on office rent levels is given in Figure 8.1.

On the other hand, keeping all other things constant, a unit increase in the RSQI of a station leads to an average increase of the rent level by 8%. The refinement of this effect is achieved by observing the effect of a change in rail service quality on the rent level at different distance ranges from the station. The estimation of the cross-effect of rail service quality with station proximity shows the effect of service quality at different distance categories. A unit increase of the rail service quality at the nearest station leads to about an 18% increase in the rent level of offices within 500 metres of a station compared with the rent levels of offices beyond 4 kilometres of a railway station. The effect is halved in areas between 500 metres and 1 kilometre. The effect of an increase in rail service quality on rent levels declines with distance from the station. A graphical illustration of the effect of rail service quality at different distance categories is given in Figure 8.2. As shown in Chapter 5, a doubling of the frequency of services on the existing network setting (which halves the average waiting time) increases the average rail service quality indices of the stations by 0.2. This increase in rail service level leads to a 3.6% increase in rent level for offices within 500 metres of a station compared with offices beyond the 4 kilometre range. In the last distance category (between 2 and 4 kilometres), the change is translated into a 1% increase in rent levels.

Similarly, a decrease in the in-vehicle time component of the generalized journey time by increasing the speed of the vehicles leads to an increase in the RSQI. For example, a 50% increase in the speed of the trains directed to the railway stations used in the analysis results on average, in an increase of the RSQI of the stations by half a unit (0.5). This in turn leads to an increase of rent by 9% for offices located within 500 metres of the stations compared with offices located beyond the 4 kilometre range. For offices located in the range of 500 metres to 1 kilometre, the effect on office rents of increasing the speed of trains by 50% is about 4.5% compared with the effect on office rents beyond 4 kilometres from a railway station. Because the RSQI of a station integrates all components of the generalized journey time, it is possible to compare the effect of changes in the time components on office space rent. Given the current setting of the railway network, doubling the frequency of train service and increasing the speed of the trains by 20 percent results in an equivalent increase of the rail service quality index and thus of office rent.

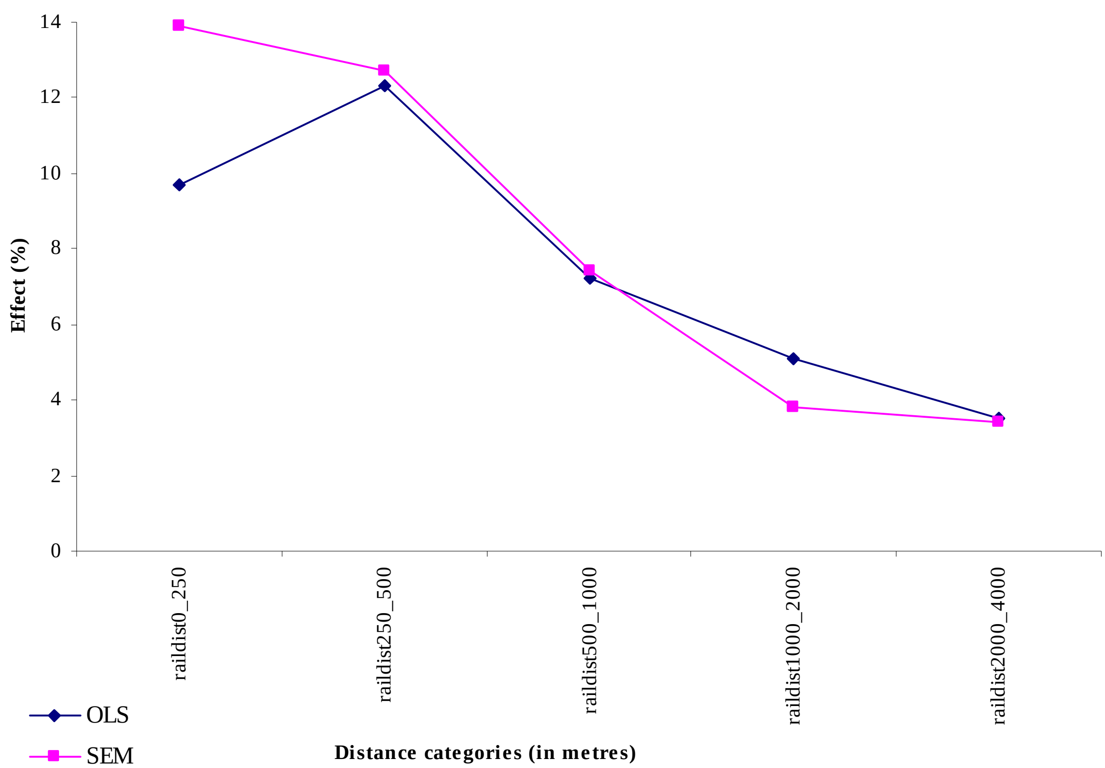


Figure 8.1: Effect of distance to nearest railway station on office rent

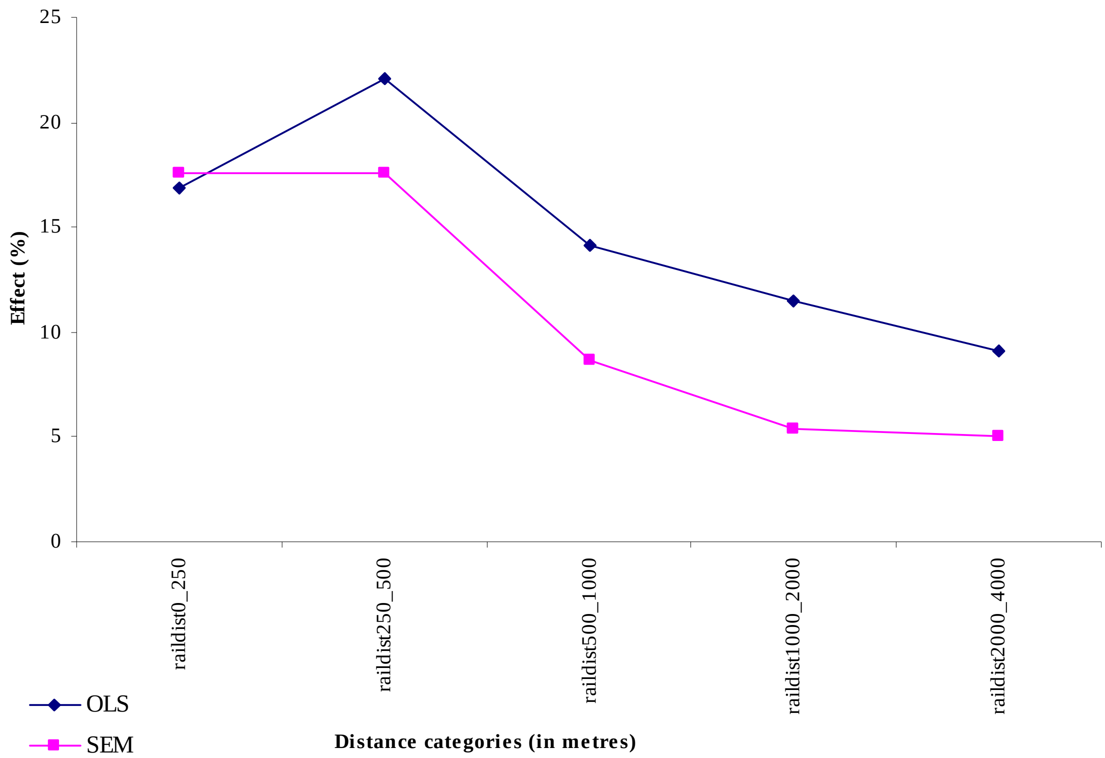


Figure 8.2: Cross-effect of RSQI and distance to the nearest station on office rent

(B) Effect of highway accessibility

Road accessibility, which includes distance to the nearest highway entry/exit point and the opportunity index defined by Equation 1, has significant effects with expected signs. The elasticities of distance to the nearest highway entry/exit and opportunity index on office rent levels are -0.044 and -0.270, respectively, for both models. This means a 1% increase in both factors leads to a decrease of 0.044% and an increase of 0.270% in office rent levels, respectively.

8.5.2 Temporal effect

Figure 8.3 below shows the temporal development of the rent prices. The rent prices can be seen to follow the development of the demand and supply of office space. The demand and supply of real estate is surveyed by Dynamis (2006). Between about 1995 and 2001, there was a relatively tight office market, which was reflected by a sharp increase in the real estate price. The shortage of office space stimulated the building of new offices, which were completed with a several-year time lag. After a peak of office floor space take-up in 2002, the demand for office floor space declined, but the supply of new offices soared as a result of the initiatives that were taken in the tight-market period. The large over-supply of office real estate after 2002 led to a decline in the real estate price. During this time, the demand for office space was also witnessed to decline. After the year 2000, there was a general demand slowdown in the office market. This is related to the general slowdown of the Dutch economy in these years.

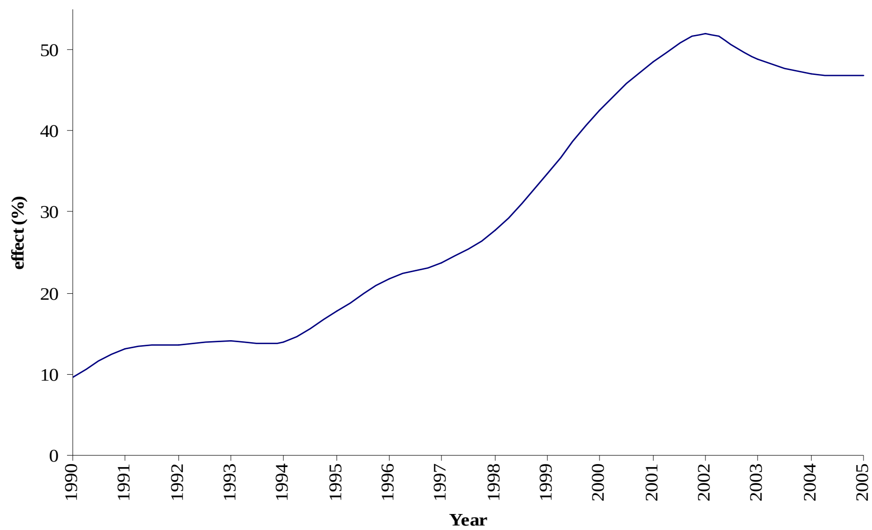


Figure 8.3: Increase in rent levels by year compared with rent levels before 1990

8.5.3 Effect of land use variable

The analysis includes several land use types. Of particular interest are the proportions of land use devoted to railway and asphalt. These are connected with the two types of accessibility factors: namely, railway and highway. They are expected to reflect the nuisance effect of both modes of transport. The nuisance effects are reflected by the negative impact of the proportion of land used for railway and highway in the postcode areas on the office rent level. However, the effects are not significant at the 10% significance level.

However, significant effects are observed from cultivation under glass, extraction of minerals, industrial areas, and waters broader than 6 metres, all factors, which have a negative impact on the office rent levels in the neighbourhood. Land use types which have a positive impact on office rent levels are forest land use, park and public gardens, dry land, and service facilities (see Table 8AI.2 in Appendix 8AI).

8.5.4 Effect of building status and business nature

The analysis found significant rent differences among offices with different building statuses. 'First-user' offices are taken as the reference group for building status. First-user offices are offices that are occupied directly after completion. Compared with this reference building status, second-user offices rent for around 11% less. On the other hand, offices occupied after renovation rent for 5% less than new offices (first-user offices). No significant difference is found for the other building statuses.

Moreover, the estimation results show that a significant rent difference is observed for some occupants' nature of business. Our analysis takes Industrial Companies and Public Utilities as a reference group. Higher rent levels are observed for offices occupied by Credit and Insurance Companies. Such offices rent for around 10% more than the rent levels of the reference group. Similarly, Financial Business Companies rent for 5% more compared with the reference group. On the other hand, offices occupied by a Trade and Repair company, and Education and Health Care rent for about 7% and 10% less than the reference group, respectively. The analysis shows no significant difference in the office rent levels of other occupant types.

Among the different types of rent contracts, we only found a significant difference between direct rent from the owner and sublease contracts. In this case, offices rented by sublease contracts are found to rent for about 5% more than direct rent contracts from the owner.

8.6 IMPLICATIONS OF SOUTH AXIS INVESTMENT FOR OFFICE RENT LEVELS

The model that has been described in the previous sections is applied to the case of the South Axis in Amsterdam (the Zuidas). At the Amsterdam South Axis, a high-status office park is being created that is meant to attract national and international head offices and other users who appreciate high quality locations (see Rienstra and Rietveld 1999). Urban development at the Amsterdam South Axis is supported by the Amsterdam South Axis project. Several alternatives have been proposed for this project, with varying quantities of real estate to be built for offices, residences, and other activities. The most ambitious alternative includes having the central rail- and motorways running through tunnels and using the space above for building more real estate. However, an explorative CBA (Besseling et al. 2003) showed none of the alternatives to have a net positive effect over the current situation with some minor adaptations to the motorway and railway station. Future developments are therefore still uncertain. Nevertheless, at the moment this area continues to be developed by building more offices. In the coming years therefore more office space will become available in the South Axis. For a description of the foreseen high speed line (HSL South), refer Section 7.6.

Effect of high speed rail on the RSQI of the South Axis station

The implementation of high speed rail reduces the in-vehicle travel time, and thereby the generalized journey time from which the RSQI used in this chapter is derived. To show the effect of an increase in the speed level of trains on the RSQI, we take the case of the South Axis Station. Figure 8.4 shows the RSQI of the station to increase, on average, at a rate of 0.1 per 10% increase in the speed of vehicles. In combination with Table 8.3, it can be seen that the rent levels for offices within 500 metres of the station increase at a rate of 1.8% for every 10% increase in the speed of trains directed to the station.

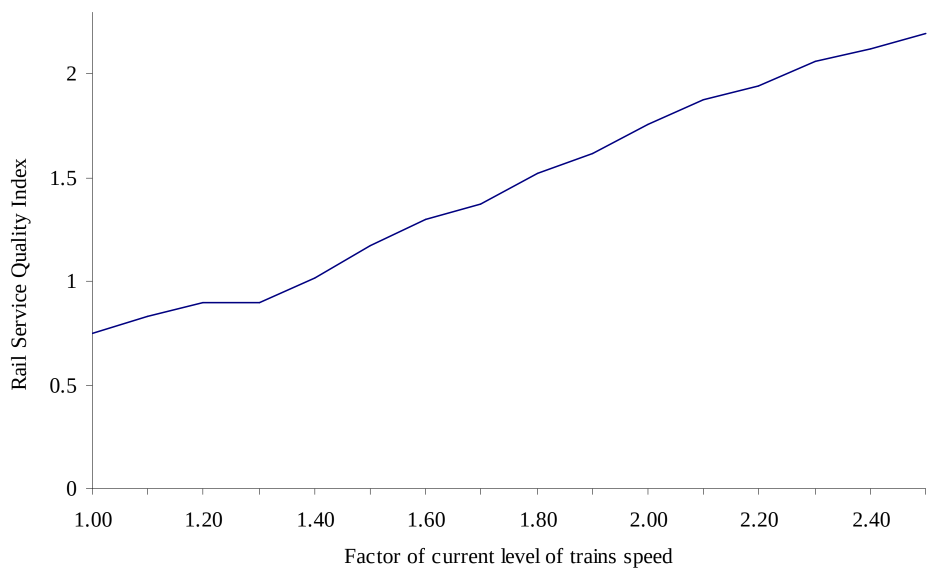


Figure 8.4: The effect of speed level (as a factor of current level of speed) of trains directed to South Axis Station on the rail service quality index (RSQI) of the station

The HSL South from Amsterdam has three national destinations: Schiphol Airport, Rotterdam and Breda. On average, the travel time is halved, implying a doubling of speed. Given the current settings, the operation of HSL South is expected to increase the RSQI of the South Axis Station by 0.30 from 0.74 to 1.04. This increase in the rail service quality index does not take into account the improvements made on the ordinary lines. Some of the ordinary lines that are directed to the South Axis Station are assumed to be terminated because the alternative routes through Breda or Rotterdam Central involve shorter trips. According to the model prediction, this increase in the RSQI at the South Axis Station leads to an increase of office rent levels within the 500 metres range of the station by about 5.4% compared with the rent of offices located beyond 4 kilometres from the station.

8.7 CONCLUSION

From the analysis of this chapter we can draw several conclusions. First, the data on office rent used in the analysis exhibits spatial autocorrelation. The appropriateness of the spatial error model for the model estimation indicates that neighbouring office locations share common features unobserved by the model. These unobserved features can range from unaccounted structural features to environmental features and latent location factors, such as the image of a site caused by the appearance of neighbouring buildings. Spatial autocorrelation models improve estimates by reducing the bias that can result due to correlation in the error components of the model.

The spatial autocorrelation model estimated in this chapter found a significant effect with expected signs for accessibility features on office rent levels. Both railway and highway accessibility are included. The main interest of this chapter is to analyse the effect of railway accessibility on office rent levels. The chapter shows the relevance of railway accessibility as measured by proximity and the RSQI for office rent in the Netherlands. Rent levels decline as the distance from the nearest railway station increases. Compared with the rents of offices located beyond 4 kilometres of the railway station, the rents of offices within 250 metres of railway station are about 14% higher. The rent difference decreases to about 7% and 4% for offices in the distance range 500 to 1000 metres and 1000 to 2000 metres, respectively, compared with offices located beyond 4 kilometres of a railway station. On the other hand, the rail service quality of a railway station has a positive effect on office rent levels. Furthermore, the cross-effect of distance and service quality on rent shows a declining effect of the RSQI of a station with distance. A stronger effect is observed on offices located in the immediate vicinity of a railway station. This shows that the range at which railway accessibility will have a meaningful effect on office rent levels is quite limited. As has been pointed out in several other earlier empirical studies, this range represents a reasonable walking distance. The meta-analysis discussion in Chapter 2 of this thesis confirms statistically that railway stations generally have a local effect on commercial property values (*see* Section 2.3.5).

The flexibility of the RSQI of a station allows us to make a model prediction based on expected changes in the railway network setting. The chapter assessed the implication of the High Speed Line (HSL) South implementation for office rent levels at the South Axis. The operation of the HSL is expected to upgrade the RSQI of the South Axis Station which, in turn, is expected to increase the rent levels of office floor space around the station. Based on the foreseen change, the chapter found that, on average, rent levels are expected to rise by 5.4% for offices located within 500 metres of the station. In reality, the effect could be somewhat higher than that for two main reasons. First, the chapter only considers the changes in HSL. Improvements in terms of the ordinary rail operation are not considered. Second, the HSL changes mainly concern changes in the national railway network. In the case of HSL operation, international origins can play a big role in upgrading the rail service quality status of the South Axis Station. However, these two aspects can easily be integrated in the model when more data are available.

APPENDIX 8AI: Spatial autocorrelation test statistics and estimation results

Spatial autocorrelation test statistics

1. Moran's *I*

The Moran's *I* test statistic is the most commonly used test for checking for spatial autocorrelation in the data. The test was developed by Moran (1948). The test statistic is specified as follows:

$I = (N / S_0)(\mathbf{e}'\mathbf{W}\mathbf{e} / \mathbf{e}'\mathbf{e})$,

where, N is the number of observations, **e** is a vector of the OLS residuals; S_0 is the standardization factor which is the sum of the elements of the weights matrix **W**. For a row standardized **W**, the Moran's *I* is reduced to ($I = \mathbf{e}'\mathbf{W}\mathbf{e} / \mathbf{e}'\mathbf{e}$). The value of the statistic ranges between -1 and 1. A value of -1 indicates perfect negative correlation, where offices with a high rent are generally neighboured by offices with lower rent, and vice versa. On the other hand, a value of 1 indicates perfect positive correlation, where offices with a high rent are neighboured by offices with a high rent, and vice versa. A value of 0 shows no spatial autocorrelation. The statistic is asymptotic to a normal distribution approximation (Cliff and Ord 1971; Sen 1976). The specifications of the Lagrange Multiplier tests (Anselin 1995) are given below

1. $LM_{\text{error}} = [\mathbf{e}'\mathbf{W}\mathbf{e} / (\mathbf{e}'\mathbf{e} / N)]^2 / [\text{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W})]$;

2. $LM_{\text{lag}} = [\mathbf{e}'\mathbf{W}\mathbf{y} / (\mathbf{e}'\mathbf{e} / N)]^2 / D$;

with $D = [\mathbf{WX\beta})'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{WX\beta}) / \sigma^2] + \text{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W})$

The diagnosis of the spatial autocorrelation is based on a series of tests:

3. Robust $LM_{\text{error}} = [\mathbf{e}'\mathbf{W}\mathbf{e} / (\mathbf{e}'\mathbf{e} / N) - T(R\tilde{J}_{\rho.\beta})^{-1}(\mathbf{e}'\mathbf{W}\mathbf{y} / (\mathbf{e}'\mathbf{e} / N))]^2 / [T - T^2 (R\tilde{J}_{\rho.\beta})^{-1}]$;
with $(R\tilde{J}_{\rho.\beta})^{-1} = [T - (\mathbf{WX\beta})'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{WX\beta}) / (\mathbf{e}'\mathbf{e} / N)]^{-1}$, and
$T = \text{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W})$;

4. Robust $LM_{\text{lag}} = [\mathbf{e}'\mathbf{W}\mathbf{y} / (\mathbf{e}'\mathbf{e} / N) - (\mathbf{e}'\mathbf{W}\mathbf{e} / (\mathbf{e}'\mathbf{e} / N))]^2 / [R\tilde{J}_{\rho.\beta} - T]$;

5. $SARMA = [\mathbf{e}'\mathbf{W}\mathbf{y} / (\mathbf{e}'\mathbf{e} / N) - \mathbf{e}'\mathbf{W}\mathbf{e} / (\mathbf{e}'\mathbf{e} / N) -] / [R\tilde{J}_{\rho.\beta} - T] + (\mathbf{e}'\mathbf{W}\mathbf{e} / (\mathbf{e}'\mathbf{e} / \mathbf{N}))^2 / T$.

All these tests are distributed as χ^2, with one degree of freedom for tests 1-4 and with two degrees of freedom for test 5.

Table 8AI. 1: Descriptive statistics of the variables included in the estimation of office rent levels

ACCESSIBILITY	N	Minimum	Maximum	Mean	Std. Deviation
Rail service quality index (RSQI-destination station)	9340	0.034	1.464	0.686	0.300
Distance to nearest railway station (metres)	11298	30	20,139	1,686	1,751
Distance to highway entry/exit points (metres)	11298	21	35,372	1,878	1,826
Opportunity index	11298	133.64	980.35	387.12	103.29
BUILDING STATUS					
First-user	1508	0	1	0.133	
Second user	9114	0	1	0.807	
Renovation	463	0	1	0.041	
Under construction	61	0	1	0.005	
Yet to be built	26	0	1	0.002	
TYPE OF CONTRACT					
Direct rent	10967	0	1	0.971	
Rent extension	109	0	1	0.010	
Sub lease	222	0	1	0.020	
TYPE OF BUSINESS					
Industrial companies and public utilities	661	0	1	0.059	
Building and civil engineering	260	0	1	0.023	
Trade and repairing companies	627	0	1	0.055	
Transportation and storage	281	0	1	0.025	
Communication companies	418	0	1	0.037	
Credit and insurance services	508	0	1	0.045	
Financial business services	685	0	1	0.061	
Other business services	2566	0	1	0.227	
Computer companies	1104	0	1	0.098	
Public administration, defence or social security	728	0	1	0.064	
Education and health care	707	0	1	0.063	
Other institutions and companies	1406	0	1	0.124	
Missing category	1347	0	1	0.119	
LAND USE					
Cultivation under glass	9357	0	0.509	0.002	0.015
Other agricultural use	9357	0	0.963	0.141	0.220
Forest	9357	0	0.708	0.028	0.083
Residential area	9357	0	0.967	0.322	0.239
Extraction of minerals	9357	0	0.119	0.001	0.005
Industrial areas	9357	0	0.941	0.104	0.169
Service facilities	9357	0	0.733	0.098	0.156
Other public facilities	9357	0	0.192	0.010	0.024
Socio-cultural facilities	9357	0	0.488	0.037	0.054
Railway	9357	0	0.443	0.024	0.037
Asphalted road	9357	0	0.277	0.054	0.039
Airport	9357	0	0.722	0.002	0.034
Park or public garden	9357	0	0.491	0.043	0.078
Sports park	9357	0	0.405	0.023	0.038
Dry natural land	9357	0	0.550	0.002	0.023
Wet natural land	9357	0	0.369	0.002	0.015
Water areas broader than 6 m	9357	0	0.582	0.047	0.056

Table 8AI. 2: *Estimation results (continuation of Table 8.3 in the main text): (t, z scores in brackets)*

Variables	Separate effect of distance and rail service quality index (RSQI)		Cross-effect of distance and rail service quality index (RSQI)	
	OLS	SEM	OLS	SEM
Second user	-0.115*** (-13.093)	-0.109*** (-12.991)	-0.115*** (-13.117)	-0.109*** (-13.009)
Renovation	-0.039** (-2.393)	-0.050*** (-3.276)	-0.039** (-2.408)	-0.050*** (-3.289)
Under construction	-0.006 (-0.132)	0.022 (0.548)	-0.010 (-0.243)	0.021 (0.521)
Yet to be built	0.026 (0.420)	0.049 (0.841)	0.024 (0.399)	0.048 (0.825)
Rent extension	0.032 (1.091)	0.024 (0.891)	0.034 (1.159)	0.026 (0.950)
Sub-lease	0.066*** (3.260)	0.048** (2.507)	0.067*** (3.294)	0.048** (2.524)
Building and civil engineering	-0.036 (-1.637)	-0.030 (-1.432)	-0.037* (-1.657)	-0.030 (-1.431)
Trade and repairing companies	-0.069*** (-4.099)	-0.067*** (-4.228)	-0.070*** (-4.157)	-0.068*** (-4.273)
Transportation and storage	-0.020 (-0.956)	-0.025 (-1.258)	-0.023 (-1.093)	-0.028 (-1.407)
Communication companies	-0.002 (-0.101)	-0.013 (-0.705)	-0.002 (-0.105)	-0.011 (-0.630)
Credit and insurance services	0.117*** (6.480)	0.107*** (6.260)	0.116*** (6.410)	0.108*** (6.313)
Financial business services	0.074*** (4.441)	0.052*** (3.304)	0.073*** (4.407)	0.052*** (3.342)
Other business services	0.018 (1.345)	0.008 (0.668)	0.018 (1.331)	0.009 (0.738)
Computer companies	-0.001 (-0.039)	-0.006 (-0.438)	-0.002 (-0.117)	-0.007 (-0.520)
Public administration, defence or social security	-0.014 (-0.845)	-0.004 (-0.271)	-0.015 (-0.877)	-0.003 (-0.197)
Education and health care	-0.108*** (-6.391)	-0.096*** (-6.063)	-0.109*** (-6.459)	-0.095*** (-6.002)
Other institutions and companies	-0.076*** (-5.255)	-0.071*** (-5.208)	-0.076*** (-5.292)	-0.070*** (-5.124)
Missing category	0.023 (1.505)	0.020 (1.410)	0.021 (1.406)	0.019 (1.358)
cultivation under glass	-1.115*** (-5.536)	-0.764*** (-3.232)	-1.061*** (-5.277)	-0.743*** (-3.147)
other agricultural use	-0.054 (-1.519)	-0.014 (-0.270)	-0.046 (-1.321)	-0.015 (-0.294)
forest	0.109** (2.227)	0.165** (2.483)	0.117** (2.393)	0.171** (2.572)

* = significant at the 10% level; ** = significant at the 5% level; *** = significant at the 1% level.

Variables (Continued)	Separate effect of distance and rail service quality index (RSQI) OLS	SEM	Cross-effect of distance and rail service quality index (RSQI) OLS	SEM
Residential area	-0.094*** (-2.705)	-0.047 (-0.958)	-0.078** (-2.252)	-0.034 (-0.697)
Extraction of minerals	-1.665*** (-2.715)	-1.471** (-2.300)	-1.650*** (-2.692)	-1.524** (-2.384)
Industrial land	-0.286*** (-7.395)	-0.175*** (-3.336)	-0.273*** (-7.076)	-0.169*** (-3.218)
Service facilities	-0.062 (-1.550)	0.144*** (2.660)	-0.056 (-1.403)	0.144*** (2.664)
Other public facilities	-0.002 (-0.013)	0.192 (1.270)	-0.050 (-0.367)	0.132 (0.869)
Socio-cultural facilities	-0.097 (-1.489)	-0.007 (-0.088)	-0.085 (-1.304)	0.012 (0.159)
Railway	-0.297*** (-3.170)	-0.132 (-1.306)	-0.300*** (-3.204)	-0.127 (-1.251)
Asphalted road	-0.108 (-1.246)	-0.073 (-0.756)	-0.106 (-1.219)	-0.068 (-0.699)
Airport	0.194** (2.089)	0.017 (0.163)	0.210** (2.257)	0.013 (0.128)
Park or public garden	0.296*** (5.657)	0.435*** (6.999)	0.296*** (5.683)	0.426*** (6.885)
Sports park	-0.167* (-1.852)	-0.052 (-0.508)	-0.139 (-1.556)	-0.055 (-0.531)
Dry natural land	0.369*** (2.763)	0.673*** (4.372)	0.436*** (3.277)	0.694*** (4.538)
Wet natural land	-0.198 (-0.987)	0.156 (0.724)	-0.177 (-0.882)	0.146 (0.679)
Water areas broader than 6 m	0.063 (0.946)	-0.156* (-1.814)	0.059 (0.890)	-0.154* (-1.793)
Year1986	-0.011 (-0.264)	-0.053 (-1.375)	-0.009 (-0.232)	-0.053 (-1.371)
Year1987	0.019 (0.493)	0.005 (0.143)	0.021 (0.536)	0.006 (0.170)
Year1988	0.026 (0.710)	-0.001 (-0.021)	0.029 (0.804)	-0.002 (-0.052)
Year1989	0.026 (0.744)	0.001 (0.034)	0.028 (0.800)	0.001 (0.026)
Year1990	0.097*** (2.803)	0.076** (2.346)	0.097*** (2.811)	0.075** (2.306)
Year1991	0.132*** (3.995)	0.112*** (3.587)	0.133*** (4.021)	0.111*** (3.555)
Year1992	0.137*** (4.123)	0.107*** (3.395)	0.137*** (4.115)	0.107*** (3.399)
Year1993	0.141*** (4.272)	0.115*** (3.694)	0.140*** (4.257)	0.114*** (3.678)

* = significant at the 10% level; ** = significant at the 5% level; *** = significant at the 1% level.

Variables (continued)	Separate effect of distance and rail service quality index (RSQI) OLS	Separate effect of distance and rail service quality index (RSQI) SEM	Cross-effect of distance and rail service quality index (RSQI) OLS	Cross-effect of distance and rail service quality index (RSQI) SEM
Year1994	0.140*** (4.344)	0.120*** (3.940)	0.140*** (4.321)	0.118*** (3.879)
Year1995	0.178*** (5.663)	0.167*** (5.597)	0.178*** (5.653)	0.166*** (5.561)
Year1996	0.217*** (6.955)	0.192*** (6.528)	0.216*** (6.932)	0.190*** (6.451)
Year1997	0.237*** (7.630)	0.220*** (7.497)	0.237*** (7.621)	0.219*** (7.449)
Year1998	0.278*** (8.910)	0.261*** (8.873)	0.279*** (8.939)	0.261*** (8.859)
Year1999	0.348*** (11.058)	0.342*** (11.504)	0.348*** (11.050)	0.342*** (11.506)
Year2000	0.426*** (13.826)	0.424*** (14.573)	0.426*** (13.811)	0.424*** (14.558)
Year2001	0.485*** (15.513)	0.493*** (16.723)	0.484*** (15.498)	0.493*** (16.706)
Year2002	0.520*** (16.742)	0.508*** (17.301)	0.520*** (16.725)	0.507*** (17.265)
Year2003	0.489 (16.128)	0.488*** (17.057)	0.489*** (16.103)	0.487*** (17.021)
Year2004	0.471*** (15.529)	0.468*** (16.336)	0.471*** (15.511)	0.468*** (16.313)
Year2005	0.468*** (15.052)	0.474*** (16.174)	0.468*** (15.055)	0.475*** (16.182)
Number of observations (N)	**9,357**	**9,357**	**9,357**	**9,357**
R-squared	**0.3602**	**0.4255**	**0.3603**	**0.4257**

* = significant at the 10% level; ** = significant at the 5% level; *** = significant at the 1% level.

Chapter 9

9 Conclusion

9.1. CONCLUSIONS

Railway stations function as nodes in transport networks and places in an urban environment. They have accessibility and environmental impacts, which contribute to property value. The literature on the effects of railway stations on property values is found to be mixed in its findings on the impact magnitude and direction, ranging from a negative to an insignificant or a positive impact. This thesis starts with attempts to explain the variation in the findings by meta-analytical procedures (*see* Chapter 2). Here we address Research Question 1 which concerns lessons to be learned from the literature. In general, the variations are attributed to the nature of data, particular spatial characteristics, temporal effects, and methodology. Railway station proximity is addressed from the perspective of two spatial considerations: a local station effect measuring the effect for properties within ¼ mile range, and a global station effect measuring the effect of coming 250 metres closer to the station. We find that the effect of railway stations on commercial property value mainly takes place at short distances. Commercial properties within the ¼ mile range are 12.2% more expensive than residential properties. Whereas the price gap between the railway station zone and the rest is about 4.2% for the average residence, it is about 16.4% for the average commercial property. At longer distances, the effect on residential property values dominates. Commuter railway stations have a consistently higher positive impact on property values compared with light and heavy railway/metro stations. The inclusion of other accessibility variables (such as highways) in the models reduces the level of reported railway station impact.

Furthermore, this thesis analyses the effect of railway investment on land prices and land use in a polycentric city under various regulatory regimes of land markets (*see* Chapter 3, which addresses Research Question 2 concerning the implications of land markets for the effect of railway investments on land prices). The introduction of a fast mode of transport (train), accessible in discrete locations, leads to an increase in city size. The stations of the “fast” mode induce dense residential settlements in their vicinity. As a result, the average residential and commercial land rents increase in both competitive and segmented land market situations, compared with the unimodal transport case. When rail investments only serve one particular centre, this leads to the growth of the advantaged centre at the expense of the other centre. An investment in the fast mode results in city growth and an increase in rent receipts. However, the effect of the investment for individual centres and their corresponding residential areas depends on the underlying land market conditions. Restrictions on commercial land use lead to increases in commercial rents, but this is more than offset by the decrease in residential land rents.

A baseline hedonic pricing model is estimated (*see* Chapter 4) to analyse the impact of railways on house prices in terms of distance to the railway station, frequency of railway services, and perpendicular distance to the railway line. Correcting for a wide range of other determinants of house prices, we find that dwellings very close to a station are on average about 25% more expensive than dwellings at a distance of 15 kilometres or more. A doubling of train frequency leads to an increase of house values of about 2.5%, ranging from 3.5% for houses close to the station to 1.3% for houses further away. Finally, we find a negative effect of distance to railways, probably due to noise effects. Two railway station references were used in the analysis: the nearest and the most frequently-chosen station in the postcode area. This distinction indicates that railway station accessibility is a more complex concept than one might think. It involves competition between railway stations. Competition between railway stations is used as a starting point for a more comprehensive analysis of railway accessibility.

The benefits of railway accessibility are concentrated at railway stations. Thus, the discussion on railway accessibility proceeds with reference to railway stations. In the literature, railway accessibility is usually measured in a rather simplistic way. This thesis introduces several methodologies on how to address railway accessibility in general and in relation to real estate in particular. A new element in this thesis is that the measurement of railway accessibility is undertaken using the estimation of spatial interaction and nested logit models (*see* Chapters 5 and 6, which address Research Questions 3 and 4 concerning the definition and operationalization of railway accessibility and the contribution of access modes for general railway accessibility). Railway accessibility is an integral function of access and station features. We found that the quality of a railway station in terms of railway service is explained well by a function incorporating the generalized journey time, the ratio of journey time to distance, and the importance level of other stations with which the station concerned has connections. It was also found that the contribution of short trips (with duration of 30 minutes or less) to railway accessibility is low. The nested logit estimation results reveal sizable contributions of access modes to general railway accessibility.

The spatial hedonic price analyses on residential price and office rent levels indicate that significant levels of spatial dependence exist in the data (*see* Chapters 7 and 8, which empirically address Research Question 5 concerning the contribution of the railway to office rents and house prices). In both cases, the spatial error model is found to be more appropriate to model the spatial dependence than the spatial lag model. The spatial model estimation for residential price found that a unit increase in the general railway accessibility measure (as defined by this thesis in Section 6.3.2) leads to a 4% price increase of residential units. However, the proximity of railway lines produce localized negative effects on house prices. Keeping other things constant, houses located within 250 metres of the railway line and houses located between 250 metres and 500 metres of the railway line sell for 5% and 2% less compared with houses located beyond 500 metres of the railway line. On the other hand, the spatial hedonic price analysis on office rent levels shows the relevance of railway accessibility as measured by proximity and the rail service quality index (RSQI), for office rent in the Netherlands. Rent levels decline as the distance from the nearest railway station increases. Compared with the rents of offices located beyond 4 kilometres of a railway station, the rents of offices within 250 metres of a railway station are about 14% higher. The rent difference decreases to about 7% and 4% for offices in the distance range 500 to 1000 metres and 1000 to 2000 metres, respectively, compared with offices located beyond 4 kilometres of a railway station. Furthermore, the cross-effect of distance and service quality on rent shows a declining effect of the rail service quality of a station with distance. A stronger effect is observed on offices located in the immediate vicinity of a railway station. This shows that the range over which railway accessibility will have a meaningful effect on office rent levels is quite limited. As has been pointed out in several other earlier empirical studies, this range represents a reasonable walking distance. The meta-analysis discussion in Chapter 2 of this thesis confirms statistically that railway stations generally have a local effect on commercial properties value (*see* Section 2.3.5). Railways produce localized negative effects on real estate values through proximity to the railway line. However, this is more pronounced in residential property value analysis. No significant effect is found on office rental levels.

The studies on the *ex-ante* effects of the High Speed Line (HSL) South in the Amsterdam South Axis have produced interesting results. These relate to research question 6 which is addressed in both Chapter 7 and Chapter 8. Based on the scheduled service of the HSL, the general railway accessibility improvement in the immediate postcode areas of the stations leads on average to an increase in the house price of about 3%. Similarly, the expected development changes in the Amsterdam South Axis with respect to the HSL South is expected to raise the rents of offices located within 500 metres of the station on average by 5.4%.

Generally proximity to railway station increases real estate price. This means that price of residential houses and rent of offices decline with distance away from a railway station. Further, there is evidence that the peak house prices and office space rents occur some distance from the station as compared to the immediate areas. This shows that railway station pose further negative effect on the immediate areas in addition to the railway noise effect which is captured by the perpendicular distance of the property to the railway line. These negative effects can be related to traffic congestion and crimes. Due to the lack of data on these areas further investigation was not carried out.

From the findings in this thesis it can be concluded that railway accessibility contributes positively towards real estate prices. However, it affects housing value and office rental levels differently. The difference in the impact stems from the accessibility orientation that dwellings and offices have towards the railway. Residential properties are generally influenced by the *departure* orientation of the railway accessibility. The trips to the railway stations relate to the *access* part of rail trips. The modal share on this part of the trip is quite uniform over bicycle, walking, public transport, and car (Rietveld 2000). Thus, railway accessibility has a wider range of influence on residential property values. On the other hand, commercial properties generally tend to be influenced by the destination orientation of railway accessibility. The trips from the railway stations to the offices represent the egress part of rail trips. The modal share of this part of the trip is dominated by walking (Rietveld 2000). Because of the limited spatial range of walking, the spatial influence of railway accessibility on commercial property value is rather limited in distance. This is generally in line with the expectation in the literature. However, the main difference of the finding of this thesis and the general literature, lays on the fact that the effect of railway accessibility on residential property values is felt for a wider range of area in the Netherlands as compared with most empirical studies originating from the US. This is attributed to first, higher modal share railway transport receives. Second, most railway station are well connected by public transport to residential area in the urban areas. The applicability of the findings of this research would suit more in an environment in which railway transport has a higher modal share and railway stations are connected by an efficient public transport network. This is a characteristic of most European cities.

9.2 RELEVANCE OF THE RESEARCH

From a scientific viewpoint this thesis presents several methodological contributions. First, it extends an existing polycentric urban model (Sivitanidou and Wheaton 1992) to a multimodality dimension. The inclusion of additional modes makes the model more realistic. Second, the thesis presents a thorough methodological approach in addressing railway accessibility. As far as the author is aware, the application of both spatial interaction models and nested logit models in addressing railway accessibility and its impact on real estate values is unique to this thesis. Third, the meta-analysis on the existing empirical studies in the area contributes to the advance in understanding the effect of railway accessibility on real estate prices. Lastly, the application of spatial autocorrelation models for the estimation of house prices and office rent levels contributes to the scarce literature in the area.

The study finds that the success of a railway development in producing the highest rent receipts depends on the underlying land market regimes for commercial and residential uses. The results can be used in any railway development project to achieve a successful outcome. Decisions on land market regimes mostly require government involvement, and this is one of the aspects that are important in the policy-making process. Similarly, it was found that the railway has different impact patterns on residential and commercial property values. In railway development projects which involve value capture schemes, different schemes can be implemented on commercial land and residential land based on the pattern of railway impact on these properties. The positive effect of railway accessibility on property value opens the potential for implementing a value capture method for co-financing investment on railway. However, the success of such a method depends on several criteria such as practicality of introduction, acceptability for various interest groups, effectiveness, potential revenue that can be generated and the operational costs (GVA Grimley 2004). Several methods of value capture are applied on real estate prices. However, assessing the value capture methods in the context of Dutch real estate market and recommending on the suitable method is beyond the scope of this thesis.

The social relevance of the findings of this thesis concerns the contribution of railway stations to the dynamics of urban areas. In order to understand this contribution one should understand the effects that the railway will have on real estate prices, since these are important signals to developers. Of particular importance is the problem of mobilizing sufficient resources for the construction of railway lines. The potential for the development of real estate around railway stations can be assessed by means of the models developed in this thesis. Hence, it is possible to find out to what extent the costs of building railway lines and railway stations can be covered by means of the participation of real estate developers. The implementation of the HSL South in the Amsterdam South Axis concerns the largest infrastructure-related urban development project in the Netherlands. Based on the accessibility projection, this study predicts the foreseen impacts on office rent and house price levels.

Another policy-relevant aspect of the research relates to the approach to the determination of general railway accessibility. It explicitly identifies the contribution of all access modes and rail service provided in a station to overall railway accessibility. The general railway accessibility level of a surrounding area, after a major investment as, for example, in the Amsterdam South Axis, can be projected. The findings can be used in any railway accessibility improvement schemes. It gives the opportunity to coordinate activities to achieve a higher accessibility level. The two possible target areas for coordination, in order to improve the general railway accessibility level, are: the service levels provided by the railway company, and the public transport service connecting the railway stations. Similar coordinated activities can be achieved between parking or park-and-ride projects and railway services. At the same time, the results of the research can be used to define the catchment area (market area) of the stations. This in turn can be used as a basis for site selection for new line development or planning extensions for existing lines, as well as parking facilities and feeder public transport operation. In addition, an understanding of the sensitivity of travellers towards the access and station features gives a station operator the basis for increasing traveller turnover.

9.3 FUTURE RESEARCH DIRECTIONS

The findings of thesis can be used as a basis for further investigations in this area. The author envisages four research areas in which the theme of this thesis can be further studied. First, based on the polycentric multimodal transport urban model, a more comprehensive urban model can be developed which includes a number of different parties such as a Railway Company, producers, households and a local authority, with an emphasis on welfare maximization in the urban economy. The interaction between the labour, land and goods markets provides a setting to assess the effect of investments in railway transport on land prices.

The second line of future research relates to the further operationalization of the railway accessibility concept. In this thesis, railway accessibility computations are based on underlying train trips which are assumed as given. This means trips by other modes are not accounted for. However, accessibility in general remains relative. The railway accessibility measures adopted in this thesis are only comparable with reference to railway stations. Cross-modal comparison is not possible. The concept of railway accessibility would acquire deeper meaning if it could be compared with accessibility provided by other modes for the main trips (e.g. car, bus). This requires the modelling of the trips made by all modes of transport. The modelling could be based on a choice analysis similar to that used in this thesis. It implies that the railway share in the total number of trips becomes endogenous.

Third, international destinations and international origins play an important role in the overall railway transport in the Netherlands. Thus, for a more refined assessment of the accessibility measure and assessment of the benefits, international trips should also be analysed together with the national railway trips.

Finally, further investigation can be done in relation to the spatial dependence analysis of real estate prices. Although the use of spatial models considerably improves the estimation outcome, the effects of the accessibility and environmental features on house prices are sensitive to the specification of the spatial models. This suggests additional investigation is required with regard to the specification of the correct spatial model.

REFERENCES:

Allen, W.B., K. Chang, D. Marchetti, and J. Pokalski (1986). "Value Capture in Transit: The Case of the Lindenwold High Speed Line". The Wharton Transportation Program, the Wharton School, University of Pennsylvania, April.

Al-Mosaind, Musaad A., Kenneth J. Dueker, and James G. Strathman (1993). "Light Rail Transit Stations and Property Values: A Hedonic Price Approach". Portland, OR: Center for Urban Studies. Preprint prepared for *Transportation Research Board* 72nd Annual Meeting, Washington D.C.

Alonso, William (1964). "Location and Land Use: Toward a General Theory of Land Rent". Cambridge, MA: Harvard University Press.

Anas, Alex and Ikki Kim (1996). "General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration". *Journal of Urban Economics* 40: 232-256.

Anas, Alex, Richard Arnott and Kennet Small (1998). "Urban Spatial Structure". *Journal of Economic Literature* 36: 1426-1464.

Anselin, L. (1988a). "Spatial Econometrics: Methods and models". Kluwer Academic Publishers, Dordrecht.

Anselin, L. (1988b). "Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity". *Geographical Analysis* 20: 1–17.

Anselin, L. (1995a). "Local Indicators of Spatial Association — LISA". *Geographical Analysis* 27: 93-115.

Anselin, Luc (1995b). "SpaceStat VERSION 1.80 USER'S GUIDE". University of Illinois, Urbana-Champaign.

Anselin, L. and R. Florax (eds) (1995) "Spatial Econometrics". Springer, Berlin.

Anselin, Luc (1999). "Spatial Econometrics". Bruton Center School of Social Sciences-University of Texas, Dallas.

Armstrong, Robert J. (1994). "Impacts of Commuter Rail Service as Reflected in Single-Family Residential Property Values". Preprint prepared for *Transportation Research Board* 73rd Annual Meeting, Washington D.C.

Armstrong, Robert J. and Daniel A. Rodríguez (2006). "An evaluation of the accessibility benefits of commuter rail in Eastern Massachusetts using spatial hedonic price function". *Transportation* 33: 21-43.

Arnott, Richard (2001). "Urban economic aggregates in monocentric and non-monocentric cities". Unpublished manuscript.

Ashford, N. and M. Bencheman (1987). "Passengers' choice of airport: an application of the multinomial logit model". *Transportation Research Record* 1147: 1-5.

Bajic, V. (1983). “The Effects of a New Subway Line on Housing Prices in Metropolitan Toronto”. *Urban Studies* 2: 147–58.

Basar, G., and C. Bhat. (2004). “A Parameterized Consideration Set Model for Airport Choice,” *Transportation Research* 38B: 889-904.

Basu, S. and Thomas G. Thibodeau (1998). “Analysis of spatial autocorrelation in house prices”. Journal of Real Estate Finance and Economics 17: 61 – 86.

Batten, D. F. (1995). “Network Cities: creative urban agglomerations for the 21th century”. *Urban Economics* 32: 313-327.

Ben-Akiva, M. (1973). Structure of Travel Passenger Demand Models, Ph. D. Dissertation, MIT.

Ben-Akiva, M. and S. Lerman (1985) “Discrete Choice Analysis: Theory and Application to Travel Demand”. MIT Press, Cambridge, MA.

Benjamin, John D. and G. Stacy Sirmans (1996). “Mass Transportation, Apartment Rent and Property Values”. *The Journal of Real Estate Research* 12(1): 1-8.

van den Berg, L. and P.M.J. Pol (1998). “The European high-speed train-network and urban development: Experiences in fourteen European urban regions”. EURICUR series. Ashgate, Aldershot

Berkhout, P.H.G. and J.P. Hop (2002). “Een prijskaartje aan kwaliteitskenmerken van kantoren”. In opdracht van DTZ Zadelhoff en Twynstra Gudde. Amsterdam, Stichting voor Economisch Onderzoek der Universiteit van Amsterdam.

Bertolini, Luca and Tejo Spit (1998). “Cities on Rails: The redevelopment of Railway Station Areas”. E & FN Spon, London.

Besseling, P., J. Ebregt, and R. Saitua, (2003). “Kengetallen kosten-batenanalyse project ‘Zuidas Amsterdam’”. CPB document no. 44. Den Haag, Centraal Planbureau.

Bollinger, C. and K. Ihlanfeldt (1997). “The impact of rapid rail transit on economic development: The case of Atlanta’s MARTA”. *Journal of Urban Economics* 42: 179-204.

Bollinger, C. and K. Ihlanfeldt (2000). “Spatial Interactions among Employment and Population Groups”. Unpublished manuscript.

Bonnafous, A. (1987). “The regional impact of the TGV”. *Transportation* 14: 127-137.

Bovy, P. H. L. and E. Stern (1990). “Route Choice: Way finding in Transport Networks”. Kluwer Academic Publishers, Boston, MA.

Bowes, David R., and Keith R. Ihlanfeldt (2001). “Identifying the Impacts of Rail Transit Stations on Residential Property Values”. *Journal of Urban Economics* 50: 1-25.

Boyce, D., B. Allen, R. Mudge, P. Slater and A. Isserman (1972). “Impact of Rapid Transit on Suburban Residential Property Values and Land Development”. Working Paper, NTIS 220693. Department of Regional Science, University of Pennsylvania.

Brigham, E. (1965). “The Determinants of Residential Land Values”. *Land Economics* 41: 325-334.

Brotchie, J. (1991). "Fast rail networks and socio-economic impacts". In: "Cities of the 21th century: new technologies and spatial systems". Edited by J. Brotchie, M. Batty, P. Hall, and J. Newton. Longman Cheshire, Melbourne, pp. 25-37.

Burkhart, R. (1976). Summary of Research: Joint Development Study. New York: the Administration and managerial Research Association of New York City.

Burridge, P. (1980). "On the Cliff-Ord test for spatial autocorrelation". *Journal of the Royal Statistical Society B* 42: 107–108.

Cambridge Systematics, Inc. (1999). "Quantifying the Impacts of a Light Rail Transit Station on Adjacent Property Values". Unpublished manuscript.

CBS (2004). http://statline.cbs.nl/StatWeb/start.asp?lp=Search/Search

Cervero, R., and B. Susantono (1999). "Rent Capitalization and Transportation Infrastructure Development in Jakarta". *Review of Urban and Regional Development Studies* 11(1): 11-23.

Cervero, Robert (1984). "Light Rail Transit and Urban Development". *Journal of the American Planning Association* 50: 133-147.

Cervero, Robert (1996). "Transit-Based Housing in the San Francisco Bay Area: Market Profiles and Rent Premiums". *Transportation Quarterly* 50(3): 33 – 47.

Cervero, Robert (1997). "Transit-Induced Accessibility and Agglomeration Benefits: A Land Market Evaluation". Institute of Regional Development, University of California, Berkeley, Working Paper 691.

Cervero, Robert and Michael Duncan (2001). "Rail Transit's Value Added: Effect of Proximity to Light and Commuter Rail Transit on Commercial Land Values in Santa Clara County California". Paper prepared for National Association of Realtors Urban Land Institute.

Cervero, Robert and Michael Duncan (2002a). "Land Value Impact of Rail Transit Services in Los Angeles County". Report prepared for National Association of Realtors Urban Land Institute.

Cervero, Robert and Michael Duncan (2002b). "Land Value Impact of Rail Transit Services in San Diego County". Report prepared for National Association of Realtors Urban Land Institute.

Chen, Hong, Anthony Rufulo and Kenneth Dueker (1998). "Measuring the Impact of Light Rail systems on Single-Family Home Prices: A Hedonic Approach with GIS Applications". Paper prepared for the *Transportation Research Board* 77th Annual Meeting, Washington D.C.

Clark, C. (1958). "Transport: maker and breaker of cities". In: " ". Piet Rietveld, Kennet Button and Peter Nijkamp (eds), *Urban Transport* (2003). Edward Elgar Publishing Limited Cheltenham, UK

Clark, William A. V. and Marrianne Kuijpers-Linde (1994). "Commuting in Restructuring Urban Regions". *Urban Studies* 31: 465-483.

Cliff, A. and J.K. Ord (1973). "Spatial Autocorrelation". London: Poin.

Cook, T.D., H. Cooper, D.S. Cordray, H. Hartmann, R.J. Light, T.A. Louis and F. Mosteller (1992). "*Meta-Analysis for Explanation: casebook*". Russell Sage Foundation, New York.

Craig, Steven G. and Pin T. Ng (2001). “Using quantile smoothing splines to identify employment subcenters in a multicentric urban area”. *Journal of Urban Economics* 49: 100-120.

Damm, D., S. R. Lerman, E. Lerner-Lam, and J. Young (1980). “Response of urban real estate values in anticipation of the Washington Metro”. *Journal of Transport Economics and Policy* 14: 315-336.

Davidson, Bill and Lidan Yang (1999). “Modeling commuter rail station choice and access mode combinations”. Paper presented at the *Transportation research Board* 78th Annual Meeting, Washington D.C.

Davoudi, S. 2002, “Polycentricity - modelling or determining reality?”. *Town and Country Planning*, April 2002: 114-117.

Debrezion, Ghebreegziabiher, Eric Pels and Piet Rietveld (2006). “The Impact of Railway Stations on Residential and Commercial Property value: A Meta-Analysis”. *Journal of Real Estate Finance and Economics*, Forthcoming.

DEGW and BCI (1998). “The impact of high speed trains on urban centres: A vision for five Dutch cities. Den Haag”. DEGW and Buck Consultants International (BCI), commissioned by Ministerie van VROM.

Dewees, D. N. (1976). “The effect of a subway on residential property values in Toronto”. *Journal of Urban Economics* 3: 357-369.

Dieleman, F.M. and A. Faludi (1998). “Randstad, Rhine-Ruhr and Flemish Diamond as one polynucleated macro-region”. *Tijdschrift voor Economische en Sociale Geografie* 89: 320-327.

Dornbusch, D. M. (1975). “BART-induced changes in property values and rents, in Land Use and Urban Development Projects, Phase I, BART: Final Report”. Working Paper WP 21-5-76, U.S. Department of Transportation and U.S. Department of Housing and Urban Development.

Dubin, Robin A. (1988). “Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms”. *The Review of Economic Studies* 70 (3): 466-474.

Dubin, Robin A. 1992. "Spatial Autocorrelation and Neighborhood Quality," *Regional Science and Urban Economics* 22: 433–452.

Dueker, Kenneth J. and Martha J. Bianco (1999). “Light Rail Transit Impacts in Portland: The First Ten Years”. Paper presented at *Transportation Research Board* 78th Annual Meeting.

Dvett, M., D. Dornbusch, M. Fajans, C. Falcke, G. Gussman, and J. Merchant (1979). “Land Use and Urban Development Project of BART: Final Report”. National Technical Information Service, Springfield, VA.

Dynamis (2006). Sprekende cijfers kantorenmarkten (2006). Dynamis, Amersfoort.

European Commission (2004). “EU Energy and Transport in Figures: statistical pocketbook 2004”

Fan, Kai-Sheng, Eric J. Miller and Daniel A. Badoe (1993). "Modeling Rail Access Mode and Station Choice". *Transportation Research Record* 1413: 49-59.

Federal Transit Administration (FTA) (2002). "Commercial property benefit of transits". Hickling Lewis Brod INC.

Fejarang, Robert A. (1994). "Impact on Property Values: A Study of the Los Angeles Metro Rail". Preprint prepared for *Transportation Research Board* 73rd Annual Meeting, Washington D. C., 9-13 January.

Ferguson, B., M. Goldberg and J. Mark (1988). "The Pre-Service Impacts of the Vancouver Light Rail Transit System on Single Family Property Values". In J. M. Clapp and S. D. Editors, *Real Estate Market Analysis: Methods and Applications*. Praeger, New York

Florax, R. and H. Folmer (1992). "Specification and estimation of spatial linear regression smodels: Monte Carlo evaluation of pre-test estimators". *Regional Science and Urban Economics* 22: 404-432.

Florax, Raymond J.G.M., Peter Nijkamp and Kenneth G. Willis (2002). "Comparative Environmental Economic Assessment". Edward Elgar Publishing Ltd., Cheltenham, UK.

Fotheringham, A.S. and M.E. O'Kelly (1989). "Spatial Interaction Models: Formulations and Applications". Kluwer Academic Publishers, Dordrecht, the Netherlands.

Fujita, Masahisa (1989). "Urban Economic theory: Land Use and City Size". Cambridge University Press, New York.

Fujita, Masahisa, Paul Krugman and Anthony J. Venables (1999). "The Spatial Economy: Cities, Regions and International Trade". The MIT Press, Cambridge, MA.

Garreau, Joel (1988). "Edge City: Life on the New Frontier". Doubleday, New York.

Gatzlaff, D. and M. Smith (1993). "The Impact of the Miami Metrorail on the Value of Residences Near Station Locations". *Land Economics* 69: 54–66.

Geary, R. (1954): "The contiguity ratio and statistical mapping", *The Incorporated* Statistician 5: 115-145.

Gillen, Kevin, Thomas Thibodeau and Susan Wachter (2001). "Anisotropic Autocorrelation in House Prices". *Journal of Real estate Finance and Economics* 23(1): 5-30.

Giuliano, Genevieve (1986). "Land-Use Impacts of Transportation Investments: Highway and Transit". In: Susan Hanson (ed.) *The Geography of Urban Transportation*. Guilford Press, New York

Giuliano, Genevieve and Kenneth A Small (1991). "Subcenters in the Los Angeles Region". *Regional Science and Urban Economics* 21: 163-182.

Glaeser, Edward L. and Matthew E. Kahn (2004). "Sprawl and Urban Growth". *Handbook of Urban and Regional Economics* Vol. 4. North Holland Press.

Goldberg, M.A. (1981). “Transportation Systems and Urban Forms: Performance Measurement and Data Requirements”. Proceedings of the International Symposium on Surface Transportation Performance, pp. 256-68.

Gordon, P. and H. W. Richardson (1989). “Gasoline Consumption and the Cities: A Reply”. *Journal of the American Planning Association* 55: 342-45.

Grass, R. G. (1992). “The estimation of residential property values around transit station sites in Washington, D.C”. *Journal of Economics and Finance* 16: 139-146.”

GVA Grimley (2004). “Developing a Methodology to Capture Land Value Uplift around Transport facilities” *Transportation Research Series.*

Hack, J. (2002). “*The Role of Transit Investment in Urban Regeneration and Spatial Development: a Review of Research and Current Practice*”. CIP Annual conference (Canada).

Haider, M. and Eric J. Miller (2000). “Effects of transportation infrastructure and location on residential land values: Application of spatial-autoregressive techniques” Transportation Research Record 1722:1-8.

Hansen, W.G. (1959). “How Accessibility Shapes Land Use”. *Journal of American Institute of Planners* 25: 73-76.

Hawke, G. R. (1971). “Railways and Economic Growth in England and Wales 1840-1870”. *The Journal of Economic History* 31(3): 718-719.

Haynes, K.E. (1997). “Labor markets and regional transportation improvements: the case of high speed trains, an introduction and review”. *The Annals of Regional Science* 31: 57-76.

Healey & Baker (1996). European real estate monitor. Healey & Baker, London.

Helsley, Robert W. and Arthur M. Sullivan (1991). “Urban Subcentre Formation”. *Regional Science and Urban Economics* 21: 255-275.

Hensher, David A., and William H. Greene (2002). “Specification and Estimation of the Nested Logit Model: Alternative Normalizations” *Transportation Research B* 36: 1-17.

Hess, Stephane and John W. Polak (2005) “Mixed Logit Modeling of Airport Choice in Multi-Airport Regions” (draft).

Higgins, Julian P.T. and Simon G. Thompson (2004). “Controlling the risk of spurious findings from meta-regression”. *Statistics in Medicine* 23: 1663-1682.

Hirota, R. (1984). “Present situation and effects of the Shinkansen”. Paper presented at the International seminar on high-speed trains, Paris, November.

Jansen, A. and D. Hanemaayer (1991). “Eisen aan de bedrijfsomgeving in de Randstad”. Den Haag, Ministerie van Economische Zaken.

Kastrenakes, C.R. (1988). “Development of a Rail Station Choice Model for NJ Transit”. *Transportation Research Record* 1162: 16-21.

van der Krabben, Edwin and Jan G. Lambooy (1993). “A Theoretical Framework for the Functioning of the Dutch Property Market”. *Urban Studies* 30: 1381-1397.

Kruk, René van der (2001) “Economic impact of wetland amenities: A special econometric analysis of the Dutch housing market”. Unpublished manuscript..

Landis, John, Robert Cervero, Subhrajit Guhathukurta, David Loutzenheiser and Ming Zhang (1995). “Rail Transit Investments, Real Estate Values, and Land Use Change: A Comparative Analysis of Five California Rail Transit Systems”. Monograph 48, Institute of Urban and Regional Studies, University of California at Berkeley.

Lee, D. B. (1973). “Case Studies and Impacts of BART on Prices of Single Family Residences”. University of California, Institute of Urban and Regional Development, Berkeley, CA.

Lewis-Workman, Steven and Daniel Brod (1997). “Measuring the Neighborhood Benefits of Rail Transit Accessibility”. *Transportation Research Record* 1576: 147-153.

Lythgoe, W. F., M. Wardman and J. P. Toner (2004) “Enhancing Rail Passenger Demand Models to Examine Station Choice and Access to the Rail Network” (draft).

Mannone, V. (1997). “Gares TGV et nouvelles dynamiques urbaines en centre ville: Le cas des villes desservies par le TGV Sud-Est”. *Les Cahiers Scientifiques du Transport* 31: 71-97.

Martellato, D., P. Nijkamp and A. Reggiani (1998). “Measurement and measures of network accessibility: economic perspectives”. In: M. Beuthe and P. Nijkamp (eds.), *Advances in European transport policy analysis,* Avebury, Aldershot.

McDonald, J. F. and C. Osuji (1995). “The effect of anticipated transportation improvement on residential land values”. *Regional Science and Urban Economics* 25: 261-278.

McDonald, John F. (1989). “Econometric Studies of Urban Population Density: A Survey”. *Journal of Urban Economics* 26: 361-385.

McMillen, Daniel P. (2001). “Polycentric Urban Structure: The Case of Milwaukee”. *Economic Perspectives* 25: 15-27.

Medda, F. (2000). “The assembled city”. Ph.D. dissertation, Vrije Universiteit, Amsterdam.

Miyao, T. (1981). *Dynamic analysis of the urban economy*. Academic Press, New York.

Modarres, Ali (2003). “Polycentricity and Transit Service”. *Transportation Research Part A* 37: 841-864.

Mohring, H. (1993). “Land rents and transport improvements: some urban parables”. *Transportation* 20: 267-283.

Moran, P.A.P. (1948). “The interpretation of statistical maps”. *Journal of the Royal Statistical Society B* 10: 243–51.

Morton, S.C., J.L. Adams, M.J. Suttorp and P.G. Shekelle (2004). “Meta-regression Approaches: What, why, when and how?” Paper prepared for: Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services.

Musterd, Sako and Ingrid van Zelm (2001). “Polycentricity, Households and the Identity of Places”. *Urban Studies* 38: 679-696.

Muth, R. F. 1969. “Cities and Housing”. University of Chicago Press, Chicago.

Nakamura, H. and T. Ueda, (1989). "The impacts of the Shinkansen on regional development". The Fifth World Conference on Transport Research, Yokohama, Vol. III, Western Periodicals, Ventura, California.

Nelson, A. C. (1992) "Effects of Elevated Heavy-Rail Transit Stations on House Prices with Respect to Neighborhood Income," *Transportation Research Record* 1359: 127-132.

Nelson, Arthur C. (1998). "Transit Stations and Commercial Property Values: Case Study with Policy and Land Use Implications". Paper presented at *Transportation Research Board* 77th Annual Meeting, Washington D.C.

NEORail II (2001). "The Effect of Rail Transit on Property Values: A Summary of Studies" (draft).

Oosterhaven, Jan and Piet Rietveld (2005). "Transport costs, location and the economy". In: Steven Brakman and Harry Garretsen (eds) "*Location and Competition*". Routledge Taylor and Francis Group, London.

Ord, J. K. and A. Getis (1995). "Local spatial autocorrelation statistics: Distributional issues and an application". *Geographical Analysis* 27: 286-306.

Ortúzar, J.D., L.G. Willumsen (2001). "Modelling Transport". John Wiley & Sons, Ltd, Chichester, UK.

Pellenbarg, P.H. (1985). "Bedrijfslokatie en ruimtelijke kognitie. Onderzoekingen naar bedrijfsverplaatsingsprocessen en de subjektieve waardering van vestigingsplaatsen door ondernemers in Nederland, Geografisch Instituut". Rijksuniversiteit Groningen, Groningen.

Pels, E., P. Nijkamp and P. Rietveld (2001). "Airport and airline choice in a multi-airport region an empirical analysis for the San Francisco bay area". *Regional Studies* 35: 1–9.

Pels, Eric, Peter Nijkamp and Piet Rietveld (2003). "Access to and competition between airports: a case study for the San Francisco Bay area". *Transportation Research Part A* 37: 71-83.

Pol, P.M.J. (2003). "The economic impact of the high-speed train on urban regions". Paper presented at the 43rd European Congress of the Regional Science Association, Jyväskylä, 27-30 August.

Quackenbush, Karl H., Erik Holst-Roness, Thomas J. Humphrey and Geof Slater (1987). "Red Line Extension to Alewife: Before/After Study - Appendices". Boston: Central Transportation Planning Staff for the Massachusetts Bay Transportation Authority.

Ricardo, David (1821) (1951) "On the principles of Political Economy and taxation". Piero Sraffa (ed.) *The works and corresponds of David Ricardo*, Vol. 1. Cambridge University Press, Cambridge, UK.

RICS, Policy Unit (2002). "Land Value and Public Transport: stage 1, Summary of findings".

Rienstra, S.A. and P. Rietveld (1999). "Spatial economic impacts of international head office locations: a case study of Amsterdam south". *European Urban and Regional Studies* 6(1): 85-89.

Rietveld, P., F. Bruinsma and D. van Vuuren (2001). "Coping with unreliability in public transport chains". *Transportation Research A* 35: 539-559.

Rietveld, Piet (2000). “The accessibility of railway stations: the role of the bicycle in the Netherlands”. *Transportation Research Part D* 5: 71-75.

Rietveld, Piet and Frank Bruinsma (1998). “Is transport infrastructure effective? Transport infrastructure and accessibility impacts on the space economy”. Advances in Spatial Science. Berlin, Springer.

Roback, Jennifer (1982). “Wage, Rent and the Quality of Life”. *The Journal of Political Economy* 90: 1257-1278.

Rosen, Sherwin (1974). “Hedonic Pricing and Implicit Markets: Product differentiation in pure competition”. *Journal of Political Economy* 82(1): 34-55.

Sands, B. (1993). “The development effects of high-speed rail stations and implications for California”. *Built Environment* 19(3/4): 257-284.

Sasaki, Komei (1989). “Transportation System Change and Urban Structure in Two Transport Mode Setting”. *Journal of Urban Economics* 25: 346-367.

Schwanen, Tim, Frans M. Dieleman and Martin Dijst (2001). “Travel Behaviour in Dutch Monocentric and Polycentric Urban Systems”. *Journal of Transport Geography* 9: 173-186.

Sen, A. (1976). “Large Sample-Size Distribution of Statistics Used in Testing for Spatial Correlation.” *Geographical Analysis* 9: 175-184.

Sen, Ashish and Tony Smith (1995). “Gravity Models of Spatial Interaction Behaviour”. Springer-Verlag, Berlin.

Sivitanidou, Rena (1996). “Do Office-Commercial Firms Value Access to Service Employment Centres? A Hedonic Value Analysis within Polycentric Los Angeles”. *Journal of Urban Economics* 40: 125-149.

Sivitanidou, Rena (1997). “Are Center Access Advantages Weakening? The Case of Office-Commercial Markets”. *Journal of Urban Economics* 42: 79-97.

Sivitanidou, Rena and William C. Wheaton (1992). “Wage and Rent Capitalization in the Commercial Real Estate Market”. *Journal of Urban Economics* 31: 206-229.

Sloterdijk, M.S. and P.J.M. Van Steen (1994). “Ruimtegebruik en ruimtelijk gedrag van ondernemingen: economisch-demografische bouwstenen”. Rijksuniversiteit Groningen, Faculteit der Ruimtelijke Wetenschappen, Groningen.

Smith, Adam (1776). “The Wealth of Nations”. Edited by Edwin Cannan, 1904. Reprint edition 1937. New York: Modern Library.

Smith, J. (2001). “Does Transit Raise Site Values Around Its Stops Enough to Pay for Itself?”. Paper at the Federal Highway Administration website.

Smith, V. Kerry and Ju-Chin Huang (1995). “Can Markets Value Air Quality? A Meta-analysis of Hedonic Property Value Models”. *Journal of Political Economy* 103(1): 209-227.

Song, Shunfeng (1996). “Some Tests of Alternative Accessibility Measures: A Population Density Approach”. *Land Economics* 72(4): 474-482.

Thompson, Simon G. and Julian P. T. Higgins (2002). “How should meta-regression analysis be undertaken and interpreted?”. *Statistics in Medicine* 21:1559-1573.

Tobler, W. (1970). “A Computer Movie Simulating Urban Growth in the Detroit Region”. *Economic Geography* 46: 234-240

Tse, Raymond Y. C. (2002). “Estimating Neighbourhood Effects in House Prices: Towards a New Hedonic Model Approach”. Urban Studies 39 (7): 1165-1180.

U.S. DOT (2005). http://www.bts.gov/publications/pocket_guide_to_transportation/2005/

Vessali, Kaveh V. (1996). “Land Use Impacts of Rapid Transit: A Review of Empirical Literature”. *Berkeley Planning Journal* 11: 71-105.

Voith, R. 1991. “Transportation, Sorting and House Values”. *Journal of the American Real Estate Urban Economics Association* 19: 117–37.

Voith, Richard (1993). “Changing Capitalization of CBD-Oriented Transportation Systems: Evidence from Philadelphia, 1970 – 1988”. *Journal of Urban Economics* 33: 361-376.

Von Thünen, Johan (1863). “Der Isolierte staat in Beziehung auf Landwirschaft and nationale konomie*”*. Munich: Pflaum.

Vrtic, M. and K.W. Axhausen (2003) “The impact of tilting trains in Switzerland: A route choice model of regional- and long distance public transport trips”. Paper presented at the *Transportation Research Board* 83rd Annual Meeting, Washington, D.C., January 2003.

Wardman, M. and G. A. Whelan (1999). “Using Geographical Information Systems to Improve Rail Demand Models”. Final report to the Engineering and Physical Science Research Council.

Weinberger, Rachel (2000). “Commercial Property Values and Proximity to Light Rail: Calculating Benefits with a Hedonic Price Model”. Paper presented at the *Transportation Research Board* 79th Annual Meeting, Washington, D.C., 9-13 January.

Weinberger, Rachel (2001). “Light Rail Proximity: Benefit or Detriment. The Case of Santa Clara County”. California. Paper presented at the *Transportation Research Board* 80th Annual Meeting, Washington, D.C., 7-11 January.

Weinstein, Bernard L. and Terry L. Clower (1999). “The Initial Economic Impact of the DART LRT system”. Unpublished manuscript.

Wheaton, William C. (1977). “Residential Decentralization, Land Rents, and the Benefits of Urban Transportation Investment”. *The American Economic Review* 67: 138-143.

Wheaton, William C. (2002). “Commuting, Ricardian Rent and House Prices Appreciation in Cities with Dispersed Employment and Mixed Land Use”. MIT Department of Economics Working Papers Series.

Wusten, H. van der and A. Faludi (1992). “The Randstad: playground for physical planners”. In: F. M. Dieleman and S. Musterd (eds) *The Randstad: A Research and policy Laboratory.*. Kluwer Academic Publishers, Dordrecht, pp. 17-38.

SAMENVATTING (SUMMARY IN DUTCH)

1. CONCLUSIES

Treinstations functioneren als knooppunten in transportnetwerken en locaties in een stedelijke omgeving. Via toegankelijkheids- en milieueffecten beïnvloeden ze de waarde van onroerend goed. Over de effecten van treinstations op de waarde van onroerend goed zijn in de literatuur uiteenlopende conclusies over de waarde en richting van het effect, variërend van waardedaling tot een niet significante of positieve waardestijging. Dit proefschrift verklaart de variatie in de bevindingen via meta-analytische procedures (zie hoofdstuk 2). Hier richten we ons op de eerste onderzoeksvraag aangaande de lessen die uit de literatuur kunnen worden geleerd. Over het algemeen worden de variaties toegeschreven aan de aard van de gegevens, in het bijzonder ruimtelijke karakteristieken, tijdelijke effecten en de methodologie. De aanwezigheid van treinstations wordt bekeken vanuit twee ruimtelijke overwegingen: een lokaal effect van het treinstation, welke het effect meet op onroerend goed binnen een straat van een kwart mijl, en een globaal effect, welke het effect meet van een verplaatsing van 250 meter in richting van het station. We vinden dat het effect van treinstations op commercieel onroerend goed voornamelijk plaats vindt op korte afstanden. Commercieel onroerend goed is binnen een kwart mijl 12,2% duurder in vergelijking met woningen. Waar het verschil in prijs tussen de zone van het treinstation en de overige zones ongeveer 4.2% voor een gemiddelde woning is, is het voor een gemiddeld commercieel pand 16.4%. Voor langere afstanden domineert het effect op de waarde van woningen. Treinstations die met name door forensen worden gebruikt hebben een consistent hogere positieve invloed op de waarde van onroerend goed vergeleken met light en heavy trein/metro station. Het opnemen van andere toegankelijkheidsvariabelen (zoals snelwegen) in de modellen vermindert het niveau van het gemiddelde effect van het treinstation.

Voorts analyseert dit proefschrift het effect van spoorweginvesteringen op landprijzen en landgebruik in een policentrische stad, waarbij diverse scenario’s met betrekking tot regulering en landmarkten worden bekeken. (zie hoofdstuk 3, waar onderzoeksvraag 2 over de implicaties van landmarkten op het effect van spoorweginvesteringen op landprijzen behandeld wordt). De introductie van een snelle wijze van transport (de trein), toegankelijk op discrete locaties, leidt tot een groei van de omvang van de stad. De stations van de “snelle” wijze van transport veroorzaken een hogere woningdichtheid in hun nabijheid. Dientengevolge stijgen de gemiddelde woningprijs en commerciële landprijzen in zowel de concurrerende als de gesegmenteerde delen van de landmarkt, vergeleken met de situatie met slechts één transport mogelijkheid. Wanneer spoorinvesteringen slechts één centrum bedienen, leidt dit tot de groei van het dit centrum ten koste van de andere centra. Een investering in het snelle vervoer resulteert in de groei van de stad en een verhoging van de huuropbrengsten. Echter, het effect van de investeringen voor individuele centra en hun overeenkomstige woongebieden hangt af van de onderliggende voorwaarden op de landmarkt. Beperkingen op commercieel landgebruik leiden tot verhogingen van commerciële huren, maar dit wordt meer dan gecompenseerd door de daling van de prijzen van land gebruikt voor woondoeleinden.

Een hedonisch prijs model wordt geschat (zie hoofdstuk 4) om het effect van spoorwegen op huisprijzen in termen van afstand tot het station, de frequentie van spoorwegdiensten en de afstand tot de rails te analyseren. Er wordt gecorrigeerd voor een groot aantal andere variabelen op huisprijzen. We vinden dat woningen vlak bij het station ongeveer 25% duurder zijn dan woningen op een afstand van 15 kilometer van het station of meer. Het verdubbelen van de treinfrequentie leidt tot een verhoging van de huiswaarden van ongeveer 2.5%, variërend van 3.5% voor huizen in de buurt van het station tot 1.3% voor huizen verder van het station vandaan. Tenslotte, vinden we een negatief effect van de nabijheid van spoorrails, waarschijnlijk ten gevolge van lawaai. In de analyse zijn twee stations opgenomen, het dichtstbijzijnde en het meest gekozen station binnen een postcodegebied. Dit onderscheid wijst erop dat de toegankelijkheid van een station complexer is dan meestal wordt aangenomen. Het impliceert dat er sprake is van concurrentie tussen stations. De concurrentie tussen stations wordt gebruikt als uitgangspunt voor een uitvoerige analyse van spoorwegbereikbaarheid.

De voordelen van spoorwegbereikbaarheid zijn geconcentreerd bij de stations. Zodoende spitst de discussie aangaande spoorwegbereikbaarheid zich toe op stations. Binnen de literatuur wordt spoorwegbereikbaarheid gemeten op een tamelijk eenvoudige manier. Dit proefschrift introduceert verscheidene methodes aangaande de analyse van spoorwegbereikbaarheid in het algemeen, en de relatie met onroerend goed in het bijzonder. Een nieuw element in dit proefschrift is dat de meting van spoorwegbereikbaarheid gebruikt maakt van de schatting van ruimtelijke interactie en geneste logit modellen. (zie hoofstuk 5 en 6 welke onderszoeksvragen 3 en 4 aangaande de definitie en uitwerking van spoorwegbereikbaarheid en de bijdrage van verschillende modaliteiten om naar het station te gaan aan spoorwegbereikbaarheid). Spoorwegbereikbaarheid is een functie van toegang- en stationseigenschappen. Wij vinden dat de kwaliteit van een treinstation in termen van de spoordienst goed kan worden verklaard door een functie van de algemene reistijd, de verhouding van reistijd en afstand en het belang van andere stations waar het station mee verbonden is. We vinden tevens dat de bijdrage van korte reizen (met een duur van 30 minuten of minder) aan spoorwegbereikbaarheid laag is. De geneste logit schattingen laten aanzienlijke bijdragen van verschillende modaliteiten om naar het station te gaan op algemene spoorwegbereikbaarheid zien.

De ruimtelijke hedonische prijsanalyses op huisprijzen en kantoorhuren laten zien dat er significante ruimtelijke afhankelijkheid in de data aanwezig is (zie hoofdstukken 7 en 8 waar onderzoeksvraag 5, aangaande de bijdrage van het spoor op kantoorhuren en huisprijzen, empirisch benaderd wordt). In beide gevallen is het spatial error model geschikter om ruimtelijke afhankelijkheid te modelleren dan het spatial lag model. De schatting van het ruimtelijk model voor de huizenprijzen laat zien dat een verhoging van één eenheid van de algemene spoorwegbereikbaarheid (zoals gedefinieerd in sectie 6.3.2 in dit proefschrift) leidt tot een verhoging van 4% in de huizenprijs. Echter de nabijheid spoorrails zorgt voor een lokaal negatief effect op de huizenprijs. Al het andere constant houdend, worden huizen gelegen binnen 250 meter van het spoor en huizen gelegen tussen de 250 en 500 meter van het spoor verkocht tegen een 5 respectievelijk 2% lagere prijs vergeleken met huizen verder gelegen dan 500 meter van het spoor. Anderzijds toont de ruimtelijke hedonische prijsanalyse de relevantie aan van spoorwegbereikbaarheid gemeten door de nabijheid en de kwaliteitsindex van de spoordienst voor kantoorhuren in Nederland. Huurprijzen dalen wanneer het dichtstbijzijnde station zich verder weg bevindt. De huurprijs van kantoren binnen 250 meter van het station ligt ongeveer 14% hoger in vergelijking met de huurprijzen van kantoren die meer dan vier kilometer zijn verwijderd van het station. Het verschil in huur neemt af tot 7 respectievelijk 4% voor kantoren tussen de 500 en 1000 meter en kantoren tussen de 1000 en 2000 meter in vergelijking met kantoren die meer dan 4 kilometer verwijderd zijn van het station. Tevens toont het kruislings effect van afstand en de kwaliteit van de dienst op de huur een dalend effect van de kwaliteit van de spoordienst op een station als de afstand toeneemt. Een sterker effect wordt voor kantoren waargenomen die in de directe nabijheid van een station worden gevestigd. Dit toont aan dat de range waarover de spoorwegbereikbaarheid een significant effect heeft op de kantoorhuren beperkt is. Zoals in vorige empirische studies is aangetoond geeft deze range een afstand die redelijker wijs te lopen is. De meta-analyse in hoofdstuk 2 bevestigt statistisch dat de treinstations over het algemeen een lokaal effect op de waard commercieel onroerend goed (zie sectie 2.3.5). De spoorwegen zorgen voor locale negatieve effecten op de waarde van onroerend goed. Dit komt vooral tot uiting in de analyse aangaande de waarde van woningen. Voor de hoogte van kantoorhuren is geen significant effect gevonden.

De studies over de ex-ante gevolgen van de hogesnelheidslijn (HSL) bij de Amsterdamse Zuidas hebben interessante resultaten opgeleverd. Deze hebben betrekking op onderzoeksvraag 6 die zowel in hoofdstuk 7 als hoofdstuk 8 worden besproken. Gebaseerd op de geplande dienstregeling van de HSL, zou de algemene verbetering van de spoorwegbereikbaarheid op het directe postcodegebied van het station resulteren in een gemiddelde verhoging van de huizenprijs van ongeveer 3%. Tevens leiden de verwachte ontwikkelingen in de Amsterdam Zuidas met respect tot HSL tot een verwachte stijging van de huren van kantoren binnen 500 meter van het station van gemiddeld 5.4%.

Over het algemeen verhoogt de nabijheid van een treinstation de prijs van onroerend goed. De betekend dat de prijs van woningen en de huur van kantoren afnemen als tot de afstand tot een treinstation toeneemt. Verder is er bewijsmateriaal dat de piek van de prijzen en huren op korte afstand liggen van het station en niet er direct naast. Dit toont aan dat het station nog andere negatieve effecten heeft naast het geluid, dat gemodelleerd is door de afstand van het kantoor tot de spoorlijn. Deze negatieve effecten kunnen betrekking hebben op verkeerscongestie en misdaad.

Naar aanleiding van de bevindingen van dit proefschrift kan worden geconcludeerd dat spoorwegbereikbaarheid positief bijdraagt aan de prijzen van onroerend goed. Er zit echter een verschil tussen de bijdrage aan huisprijzen en kantoorhuren. Het verschil wordt veroorzaakt door de toegankelijkheid die woningen en kantoren hebben richting het spoor. Woningen worden in het algemeen beïnvloed door de vertrekmogelijkheden van de spoorwegbereikbaarheid. De reizen naar het treinstation zijn gerelateerd aan het toegangsdeel van de treinreis. Het modale aandeel op dit deel van de reis is vrij uniform over de fiets, lopen, openbaar vervoer en de auto (Rietveld, 2000). Dus, heeft de spoorwegbereikbaarheid een grotere invloed op de waarde van woningen. Aan de andere kant worden commerciële panden meer beïnvloed door de aankomstmogelijkheden van de spoorwegbereikbaarheid. De trips van de treinstations naar de kantoren vormen het laatste deel van de totale reis. Lopen is de meest gekozen modaliteit op dit deel van de reis (Rietveld, 2000). Wegens de beperkte afstand die wordt gelopen is het ruimtelijke effect van spoorwegbereikbaarheid op de waarde van commercieel onroerend goed beperkt in afstand. Dit stemt over het algemeen overeen met de verwachtingen binnen de literatuur. Echter, het voornaamste verschil van de bevindingen van dit proefschrift is dat het effect van spoorwegbereikbaarheid op de waarde van woningen voor een groter gebied in Nederland geldt in vergelijking met de meeste empirische studies over de Verenigde Staten. Ten eerste kan dit worden toegeschreven aan het hogere modale aandeel van vervoer per trein. Ten tweede, zijn de meeste stations goed verbonden met openbaar vervoer naar woongebieden in de steden. De toepasselijkheid van de bevindingen van dit onderzoek past meer in een omgeving waarin het spoorwegvervoer een hoger modaal aandeel heeft en de treinstations verbonden zijn met een efficiënt openbaar vervoer netwerk. Dit is een kenmerk van de meeste Europese steden.

2. RELEVANTIE VAN HET ONDERZOEK

Dit proefschrift heeft verscheidene wetenschappelijke bijdragen op methodologisch gebied. Ten eerste breidt het een bestaand polycentrisch stedelijk model (Sivitanidou en Wheaton, 1992) uit door meerdere modaliteiten op te nemen. Dit het model realistischer. Ten tweede bevat het proefschrift een grondige methodologische benadering van spoorwegbereikbaarheid. Voor zover de auteur weet, is de toepassing van zowel de beide ruimtelijke interactie modellen alsmede geneste logit modellen voor spoorwegbereikbaarheid en zijn impact op de onroerend goed waarde uniek. Ten derde, draagt de meta-analyse van bestaand empirisch onderzoek bij tot een beter begrip van spoorwegbereikbaarheid op de prijzen van onroerend goed. Tenslotte draagt de toepassing van ruimtelijke autocorrelatie modellen de huisprijzen en kantoorhuren bij aan de beperkte hoeveelheid literatuur in het gebied.

De studie toont aan dat het succes van spoorwegontwikkeling in het genereren van een zo hoog mogelijke huuropbrengst afhangt van de onderliggende landmarkt en beleid betreffende commercieel en woongebruik. De resultaten kunnen in elk spoorwegontwikkelingsproject worden gebruikt om een succesvol resultaat te bereiken. Besluiten over de structuur van de landmarkt vereisen meestal betrokkenheid van de overheid; dit is één van de aspecten die van belang zijn in het maken van beleid. Op een zelfde manier kwam naar voren dat het spoor de waarde van woningen dan wel de waarde van commercieel onroerend goed op verschillende manieren beïnvloedt. Bij spoorwegontwikkelingsprojecten waar men te maken heeft met een value capture regeling, kunnen verschillende regelingen worden geïmplementeerd voor commercieel land en voor woonlocaties, welke gebaseerd zijn op de manier waarop de waarde van het land beïnvloed wordt door de ontwikkelingen. Het positieve effect van spoorwegbereikbaarheid op onroerend goed opent de mogelijkheid van een value capture methode voor medefinanciering van spoorweginvesteringen. Het succes van een dergelijke methode hangt van verscheidene criteria af, zoals het praktische aspect van de introductie, aanvaardbaarheid voor diverse belangengroepen, doeltreffendheid, de potentiële opbrengst en de operationele kosten (GVA Grimley, 2004). Verscheidene methoden van value capture worden toegepast op de prijzen van onroerend goed. De beoordeling van de value capture methode in de context van de Nederlandse onroerend goed markt en het adviseren van de geschikte methode ligt echter buiten de scope van dit proefschrift.

De sociale relevantie van de bevindingen van dit proefschrift betreft de bijdrage van treinstations op de dynamica van stedelijke gebieden. Om deze bijdrage te begrijpen moet men de invloeden die spoorwegen hebben op de onroerend goed prijzen begrijpen, aangezien dit belangrijke signalen aan ontwikkelaars zijn. Van bijzonder belang is het probleem om voldoende middelen te verkrijgen voor de bouw van spoorwegen. Het potentieel voor de ontwikkeling van onroerend goed rond treinstations kan door middel de modellen, die in dit proefschrift worden ontwikkeld, worden beoordeeld. Zodoende is het mogelijk om te weten te komen in welke mate de kosten van het bouwen van spoorlijnen en treinstations kunnen worden gedekt door de participatie van vastgoedontwikkelaars. De implementatie van de HSL-Zuid in de Amsterdamse Zuidas betreft het grootste op infrastructuur betrekking hebbende stedelijke ontwikkelinsproject in Nederland. Gebaseerd op de toegankelijkheidsprojectie, voorspelt deze studie de voorziene effecten op kantoorhuren en de huisprijzen.

Een ander beleidsrelevant aspect van het onderzoek heeft betrekking op de benadering van de algemene spoorwegbereikbaarheid. Het identificeert expliciet de bijdrage aan de algehele spoorwegbereikbaarheid van alle vervoersmogelijkheden naar het station en de spoordiensten die verleend worden op een station. Het algemene niveau van de spoorwegbereikbaarheid van een omringend gebied, na een belangrijke investering (bijvoorbeeld de Amsterdam Zuidas), kan worden bepaald. De resultaten kunnen voor elk spoorwegontwikkelingsproject worden gebruik. Het geeft de mogelijkheid om activiteiten te coördineren om een hoger toegankelijkheidsniveau te bereiken. De twee mogelijke doelgebieden voor coördinatie, om het algemene niveau van de spoorwegbereikbaarheid te verbeteren, zijn: het niveau van de diensten geboden door het spoorwegbedrijf en de openbaar vervoerdiensten die de treinstations verbind. Een zelfde soort coördinatie kan worden bereikt tussen parkeer of park-and-ride projecten en spoorwegdiensten. Tevens kunnen de resultaten van het onderzoek worden gebruikt om het marktgebied van de treinstations te bepalen. Dit kan dan worden gebruikt als een basis voor de plaatsselectie voor de ontwikkeling van een nieuwe lijn of de planning van uitbreidingen van bestaande lijnen, evenals parkeerfaciliteiten en ontsluitend openbaar vervoer. Bovendien geeft het begrip van de gevoeligheid van reizigers naar toegang en stationeigenschappen een stationexploitant de basis om de omzet te verhogen.

3. MOGELIJKHEDEN VOOR VERVOLGONDERZOEK

De bevindingen van dit proefschrift kunnen als basis dienen voor verder onderzoek op dit gebied. Er zijn vier onderzoekgebieden waarop het onderwerp van dit proefschrift verder kan worden onderzocht. Ten eerste, kan een uitgebreider stedelijk model met een nadruk op welvaartsmaximalisering in de stedelijke economie worden ontwikkeld, gebaseerd op het stedelijke model voor polycentrisch multimodaal transport. Dit model omvat een aantal verschillende partijen, zoals de spoorwegen, producenten, huishoudens en een locale autoriteit. De interactie tussen de markten voor arbeid, land en goederen verschaft een speelveld waarin het effect van investeringen in spoorwegvervoer op landprijzen kan worden beoordeeld.

Het tweede toekomstige onderzoeksgebied heeft betrekking op het verdere operationaliseren van het concept van toegankelijkheid tot spoorwegen. In dit proefschrift zijn berekeningen van spoorwegbereikbaarheid gebaseerd op onderliggende treinreizen, die als gegeven worden verondersteld. Dit betekent dat geen rekening wordt gehouden met reizen die door andere modaliteiten worden uitgevoerd. Echter, toegankelijkheid in het algemeen blijft een relatief begrip. De maatstaven voor toegankelijkheid tot spoorwegstations die gebruikt zijn in dit proefschrift, zijn alleen vergelijkbaar met betrekking tot spoorwegstations. Een vergelijking tussen modaliteiten is niet mogelijk. Het concept van spoorwegbereikbaarheid zou een grotere betekenis hebben als het kon worden vergeleken met toegankelijkheid verschaft door andere modaliteiten voor de belangrijkste reizen (bijvoorbeeld auto of bus). Dit vereist het modelleren van de reizen die gemaakt zijn door alle transportmodaliteiten. Het modelleren zou op een keuzeanalyse kunnen worden gebaseerd die vergelijkbaar is met die in dit proefschrift. Dit impliceert dat het spoorwegaandeel in het totale aantal reizen endogeen wordt.

Ten derde, spelen internationale herkomsten en bestemmingen een belangrijke rol in het totale spoorwegvervoer in Nederland. Dus, voor een meer verfijnde beoordeling van de maatstaf voor toegankelijkheid en de beoordeling van de voordelen, zouden de internationale reizen per spoor samen met de nationale reizen moeten worden geanalyseerd.

Een laatste richting voor verder onderzoek is gerelateerd aan de analyse van ruimtelijke afhankelijkheid van vastgoed prijzen. Hoewel het gebruik van ruimtelijke modellen de schattingsuitkomsten aanzienlijk verbetert, zijn de effecten van de toegankelijkheids- en omgevingseigenschappen op huizenprijzen gevoelig voor de specificatie van de ruimtelijke modellen. Dit veronderstelt dat extra onderzoek is vereist met betrekking tot de specificatie van het correcte ruimtelijke model.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

339. K. STAAL, *Voting, public goods and violence.*
340. R.H.J. MOSCH, *The economic effects of trust. Theory and empirical evidence.*
341. F. ESCHENBACH, *The impact of banks and asset markets on economic growth and fiscal stability.*
342. D. LI, *On extreme value approximation to tails of distribution functions.*
343. S. VAN DER HOOG, *Micro-economic disequilibrium dynamics.*
344. B. BRYS, *Tax-arbitrage in the Netherlands evaluation of the capital income tax reform of January 1, 2001.*
345. V. PRUZHANSKY, *Topics in game theory.*
346. P.D.M.L. CARDOSO, *The future of old-age pensions: Its implosion and explosion.*
347. C.J.H. BOSSINK, *To go or not to go…? International relocation willingness of dual-career couples.*
348. R.D. VAN OEST, *Essays on quantitative marketing models and Monte Carlo integration methods.*
349. H.A. ROJAS-ROMAGOSA, *Essays on trade and equity.*
350. A.J. VAN STEL, *Entrepreneurship and economic growth: Some empirical studies.*
351. R. ANGLINGKUSUMO, *Preparatory studies for inflation targeting in post crisis Indonesia.*
352. A. GALEOTTI, *On social and economic networks.*
353. Y.C. CHEUNG, *Essays on European bond markets.*
354. A. ULE, *Exclusion and cooperation in networks.*
355. I.S. SCHINDELE, *Three essays on venture capital contracting*.
356. C.M. VAN DER HEIDE, *An economic analysis of nature policy.*
357. Y. HU, *Essays on labour economics: Empirical studies on wage differentials across categories of working hours, employment contracts, gender and cohorts.*
358. S. LONGHI, *Open regional labour markets and socio-economic developments: Studies on adjustment and spatial interaction.*
359. K.J. BENIERS, *The quality of political decision making: Information and motivation.*
360. R.J.A. LAEVEN, *Essays on risk measures and stochastic dependence: With applications to insurance and finance.*

361. N. VAN HOREN, *Economic effects of financial integration for developing countries.*
362. J.J.A. KAMPHORST, *Networks and learning*.
363. E. PORRAS MUSALEM, *Inventory theory in practice: Joint replenishments and spare parts control.*
364. M. ABREU, *Spatial determinants of economic growth and technology diffusion.*
365. S.M. BAJDECHI-RAITA, *The risk of investment in human capital*.
366. A.P.C. VAN DER PLOEG, *Stochastic volatility and the pricing of financial derivatives.*
367. R. VAN DER KRUK, *Hedonic valuation of Dutch Wetlands*.
368. P. WRASAI, *Agency problems in political decision making.*
369. B.K. BIERUT, *Essays on the making and implementation of monetary policy decisions*.
370. E. REUBEN, *Fairness in the lab: The effects of norm enforcement in economic decisions.*
371. G.J.M. LINDERS, *Intangible barriers to trade: The impact of institutions, culture, and distance on patterns of trade.*
372. A. HOPFENSITZ, *The role of affect in reciprocity and risk taking: Experimental studies of economic behavior.*
373. R.A. SPARROW, *Health, education and economic crisis: Protecting the poor in Indonesia.*
374. M.J. KOETSE, *Determinants of investment behaviour: Methods and applications of meta-analysis.*
375. G. MÜLLER, *On the role of personality traits and social skills in adult economic attainment*.
376. E.H.B. FEIJEN, *The influence of powerful firms on financial markets.*
377. J.W. GROSSER, *Voting in the laboratory.*
378. M.R.E. BRONS, M*eta-analytical studies in transport economics: Methodology and applications*.
379. L.F. HOOGERHEIDE, *Essays on neural network sampling methods and instrumental variables.*
380. M. DE GRAAF-ZIJL, *Economic and social consequences of temporary employment.*
381. O.A.C. VAN HEMERT, *Dynamic investor decisions.*
382. Z. ŠAŠOVOVÁ, *Liking and disliking: The dynamic effects of social networks during a large-scale information system implementation.*
383. P. RODENBURG, *The construction of instruments for measuring unemployment.*
384. M.J. VAN DER LEIJ, *The economics of networks: Theory and empirics.*
385. R. VAN DER NOLL, *Essays on internet and information economics.*
386. V. PANCHENKO; *Nonparametric Methods in Economics and Finance: Dependence, Causality and Prediction.*

387. C.A.S.P. SÁ, *Higher education choice in The Netherlands: The economics of where to go.*

388. J. DELFGAAUW, *Wonderful and woeful work: Incentives, selection, turnover, and workers' motivation.*